Final Report Airforce Grant #AORD-08-4077

Michael Barnsley made two visits to Wright Patterson to discuss opportunities for deployment of superfractals. Based on collaborative discussions with Dr. Ewing's team and Dan Repperger, and feedback from the presentation made during the October 2008 trip, he made a second presentation in March 2009, at "Radar Workshop on SuperResolution". In this lecture he argued that it is feasible to use superfractals in the design of antennas. This would have advantages over standard fractal designs: (i) implementation of V-variable models, whereby the structure of the antenna is of several characters at each level; (ii) simplicity of reconfiguration to different designs, for testing. The approach would be based on the use of fractal homeomorphisms.

A number of papers related to this topic were written during the period of the grant and are listed below. The published versions of [1] and [2] should include explicit acknowledgements of support from AOARD. (Barnsley also published other papers during this period but these were not related to the topic.)

A high level programmer with Matlab expertise was retained to investigate the feasibility of making an image recognition system based on fractal homeomorphisms, in a Matlab environment. In view of the decision to focus on antenna design, the poor performance of Matlab code in comparision with Barnsley's existing code, and the expense of continuing to work with Matlab, this direction was discontinued.

Some progress towards an intellectually satisfying model for living systems, using discrete superfractals, was made.

References

[1] Ross Atkins, M. F. Barnsley, David C. Wilson, Andrew Vince, A characterization of point-fibred affine iterated function systems, submitted for publication, (2009) *arXiv:0908.1416v1.*

[2] M. F. Barnsley, The life and survival of mathematical ideas, *Notices Am. Math. to appear, January 2010.*

[3] ———, Transformations between fractals, *Progress in Probability,* **61** (2009) 227-250.

[4] M. F. Barnsley, J. Hutchinson, Ö. Stenflo, V-variable fractals: fractals with partial self similarity, *Advances in Mathematics*, **218** (2008) 2051-2088.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE
11 FEB 2010

2. REPORT TYPE
Final

3. DATES COVERED
15-05-2008 to 15-08-2009

4. TITLE AND SUBTITLE
Fractal Image Filters for Specialized Image Recognition Tasks

5a. CONTRACT NUMBER
FA48690814077

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Michael Barnsley; Ross Atkins; Andrew Vince; David Wilson

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Australia National ,University,John Dedman Building #27, Australia National University,Canberra, Australia,AU,0200

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Asian Office of Aerospace Research & Development, (AOARD), Unit 45002, APO, AP, 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AOARD

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AOARD-084077

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Michael Barnsley made two visits to Wright Patterson to discuss opportunities for deployment of superfractals. In these lectures he argued that it is feasible to use superfractals in the design of antennas. This would have advantages over standard fractal designs: (i) implementation of V -variable models, whereby the structure of the antenna is of several characters at each level; (ii) simplicity of reconfiguration to different designs, for testing. The approach would be based on the use of fractal homeomorphisms. A number of papers related to this topic were written during the period of the grant. A high level programmer with Matlab expertise was retained to investigate the feasibility of making an image recognition system based on fractal homeomorphisms, in a Matlab environment. In view of the decision to focus on antenna design, the poor performance of Matlab code in comparision with Barnsley.s existing code, and the expense of continuing to work with Matlab, this direction was discontinued. Some progress towards an intellectually satisfying model for living systems, using discrete superfractals, was made.

15. SUBJECT TERMS
Vision, computer vision, Information Fusion, fractal homeomorphisms, superfractals

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**42**	

THE LIFE AND SURVIVAL OF MATHEMATICAL IDEAS

MICHAEL F. BARNSLEY

Nature and evolution provide the notion of a creative system: a core stable form (DNA), a fertile environment, a determination to survive, and random stimuli. Analogously, the mind of a mathematician provides a locus for creative systems, a place where mathematical structures live and evolve.

According to the King James version of *Genesis,* "On the Fifth day ... God created great whales." But Darwin went one better; at the end of his masterwork, in simple beautiful language, he proposes that what was created was a creative system. The last paragraph of *Origin of Species* says:

> It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us. .... There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone circling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved.

Faced with the extraordinary richness and complexity of the physical observable universe, of which we are part, what on earth can a mathematician truly create? The answer is: a vast landscape of lovely constructions, born for the first time, to live on in the realm of ideas. For the realm of ideas belongs to sentient beings such as us: whether or not there was a Creator, it is certain that the system of which we are part is, *by its very nature*, creative. Our genes are creative, they have to be, and they have to allow creative mutations. They must be stable in their creativity. Their creativity is the well-spring of ours. Not only must our genes, through the mechanisms of biology and random mutation, invent new viable forms: they must be prone to do so.

Our mathematical creativity may actually be initiated by random events at the deepest level, after deductive reasoning, consistencies, experience, and even intuition, are factored out of the process. But the creative mind contains something much more important than a random idea generator; it provides an environment in which the wild seed of a new idea is given a chance to survive. It is a fertile place. It has its refugias and extinctions. In giving credit for creativity we really praise not random generation but the determination to give life to new forms.

But when does a newly thought up mathematical concept, $\mathcal{C}$, survive? Obviously, $\mathcal{C}$ must be consistent with mathematics, true, correct, etc. But I believe that what causes $\mathcal{C}$ to survive in the minds and words of mathematicians is that it is, itself,

a *creative system.* For now I will resist trying to give a precise definition: for this young notion to survive, it needs to be adaptable. Roughly, I mean that $\mathcal{C}$ has the following attributes: (I) $\mathcal{C}$ is able to define diverse forms and structures, call them *plants*; (II) *plants* possess DNA; (III) $\mathcal{C}$ is stable in several senses; (IV) $\mathcal{C}$ can be treated from diverse mathematical positions (e.g. topological, geometrical, measure theoretic, algebraic); (V) $\mathcal{C}$ is highly adaptable and can be translated into the languages of various branches of science and engineering, with real applications; (VI) creative systems beget new creative systems.

For more that two thousand years the key forms of Euclid's *Geometry* have survived, shifted in importance, and evolved. It has all six properties of a creative system; indeed one could find a number of different ways of defining it as such. Here is the one that I like: the diverse forms and structures are objects such as lines, circles and other constructions; the DNA of these plants are formulas, such as "the equation for a straight line", provided by Descartes analytic geometry; Euclidean geometry can be treated from geometric, algebraic and other viewpoints; the topic is stable both in the sense that nearby DNA yields nearby forms and in the sense that small changes in the axioms lead to new viable geometries; it has adapted to many branches of science and engineering, with rich applications; and Euclidean geometry begat projective geometry via the inclusion of the line at infinity. Alternatively one might describe Euclid's geometry more abstractly so that the theorems are its the diverse structures and the axioms and definitions are its DNA.

Dynamical systems [19] and *cellular automata* [35] provide two recent examples of creative systems. I mention these topics because each has an obvious visible public aspect, more colorful than lines and circles drawn on papyrus: their depth and beauty are advertised to a broad audience via computer graphics representations of some of their flora. They are alive and well, not only in the minds of mathematicians, but also in many applications.

In your own mind you give local habitation and a name to some special parts of mathematics, your creative system. Since this note is a personal essay, I focus on ideas extracted from my own experience and research. In particular, I discuss *iterated function systems* as a creative system, to illustrate connections with artistic creativity. To sharpen the presentation I focus almost exclusively on point-set topology aspects. While the specifics of iterated function systems may not be familiar to you, I am sure that the mathematical framework is similar to ones that you know.

1. Iterated function systems, their attractors, and their DNA

An *iterated function system (IFS)*,

$$\mathcal{F} := (\mathbb{X}; f_1, ..., f_N),$$

consists of a complete metric space $\mathbb{X}$ together with a finite sequence of continuous functions, $\{f_n : \mathbb{X} \to \mathbb{X}\}_{n=1}^{N}$. We say that $\mathcal{F}$ is a *contractive IFS* when all its functions are contractions. A typical *IFS creative system* may consist of all IFSs whose functions belong to a restricted family, such as affine or bilinear transformations acting on $\mathbb{R}^2$.

Let $\mathbb{H}$ denote the set of nonempty compact subsets of $\mathbb{X}$. We equip $\mathbb{H}$ with the Hausdorff metric, so that it is a complete metric space. The Hausdorff distance between two points in $\mathbb{H}$ is the least radius such that either set, dilated by this

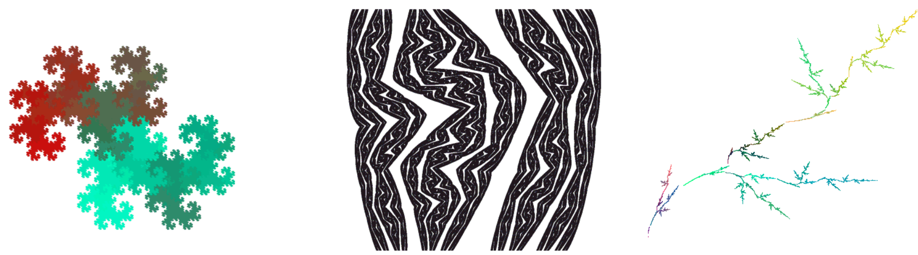


FIGURE 1. Pictures of attractors of affine, bilinear, and projective IFSs.

radius, contains the other set. We define a continuous mapping $\mathcal{F} : \mathbb{H} \to \mathbb{H}$ by

$$\mathcal{F}(B) = \cup f_n(B)$$

for all $B \in \mathbb{H}$. Note that we use the same symbol $\mathcal{F}$ for the IFS and for the mapping.

Let us write $\mathcal{F}^{\circ k}$ to denote the composition of $\mathcal{F}$ with itself k times. Then we say that a set $A \subset \mathbb{X}$ is an *attractor* of the IFS $\mathcal{F}$ when $A \in \mathbb{H}$ and there is an open neighborhood $\mathcal{N}$ of A such that

$$\lim_{k\to\infty} \mathcal{F}^{\circ k}(B) = A$$

for all $B \subset \mathcal{N}$ with $B \in \mathbb{H}$. Since $\mathcal{F} : \mathbb{H} \to \mathbb{H}$ is continuous we have $\mathcal{F}(A) = A$. Notice that our definition of attractor is topological: in the language of dynamical systems, A is a strongly stable attractive fixed-point of $\mathcal{F}$.

Our first theorem provides a sufficient condition for an IFS to possess an attractor, one of the "plants of many kinds" of an IFS creative system.

Theorem 1. [18] *Let $\mathcal{F} = (\mathbb{X}; f_1, ..., f_N)$ be a contractive IFS. Then $\mathcal{F} : \mathbb{H} \to \mathbb{H}$ is a contraction, and hence, by Banach's contraction theorem, $\mathcal{F}$ possesses a unique global attractor.*

Attractors of IFSs are our main examples of the diverse forms and structures of an IFS creative system, see Figure 1, as in attribute (I). An *affine* IFS is one in which the mappings are affine on a Euclidean space. Attractors of affine IFSs are the bread-and-butter sets of fractal geometers, such as Sierpinski triangles, twin-dragons, Koch curves, Cantor sets, fractal ferns, and so on. The geometries and topologies of these attractors are so rich, fascinating and diverse that deep papers get written about a single species, or very small families of them!

We define the *DNA* of an IFS attractor to be an explicit formula for the IFS. We refer to the DNA of an IFS attractor as an *IFS code*. The DNA for the canonical Cantor set is $(\mathbb{R}; f_1(x) = x/3, f_2(x) = (x+2)/3)$; these few symbols and their context define a non-denumerable set of Lebesque measure zero. Similarly, the DNA for the Sierpinski triangle is $(\mathbb{R}^2; f_1(x,y) = (x/2, y/2),\ f_2(x,y) = (x/2+1/2, y/2),\ f_3(x,y) = (x/2, y/2+1/2))$. Here a curve whose points are all branch points is captured in a short strand of symbols. Other simple IFS codes provide DNA for classical objects such as arcs of parabolas, line segments, triangles, and circles.

How do the individual numbers in IFS codes relate to the properties of the attractors that they define? Similarly, we might ask about the relationship between the DNA of a biological plant and the plant itself, the details of its leaf shapes, the structure of its vascular bundles, and so on.

2. When does an affine IFS possess an attractor?

In discussing this seemingly simple question we reveal how IFS theory is subtle and leads into applications, as in attribute (V). We characterize both geometrically and metrically, see attribute (IV), those affine IFSs that possess attribute (I).

Intuition incorrectly suggests that the answer to our question is: if the magnitudes of all of the eigenvalues of the linear parts of the maps of the IFS are less than one, then the affine IFS has an attractor. The situation seems to be analogous to the situation for discrete dynamical systems, [17] Proposition, p.279, where an affine map has an attractive fixed point if and only if the norm of the linear part is less than one. But the situation is not analogous. Consider for example the IFS

$$(\mathbb{R}^2; f_1(x,y) = (2y, -x/3), f_2(x,y) = (-y/3, 2x)).$$

The point $O = (0,0)$ is an attractive fixed point for both f_1 and f_2, because $f_1^{\circ 2n}(x,y) = f_2^{\circ 2n}(x,y) = (-2/3)^n (x,y)$, $f_1^{\circ(2n+1)}(x,y) = (-2/3)^n f_1(x,y)$, and $f_2^{\circ(2n+1)}(x,y) = (-2/3)^n f_2(x,y)$ for all points (x,y) in $\mathbb{R}^2$. But $\{O\}$ is not an attractor for the IFS because $(f_1 \circ f_2)^n(x,y) = (4^n x, y/9^n)$ implies that O is an unstable fixed point.

The following theorem contains an answer to our question, and an affine IFS version of the converse to Banach's contraction theorem [13]. For us, most importantly, it provides both a metric and a geometrical characterization of viable affine IFS codes. See also Berger and Wang, [9], and Daubechies and Lagarias, [10].

Theorem 2. [1] *If $\mathcal{F} = (\mathbb{R}^M; f_1, ..., f_N)$ is an affine IFS then the following statements are equivalent.*

(1) $\mathcal{F}$ possesses an attractor.

(2) There is a metric, Lipshitz equivalent to the Euclidean metric, with respect to which each f_n is a contraction.

(3) There is a closed bounded set $K \subset \mathbb{R}^M$, whose affine hull is $\mathbb{R}^M$, such that $\mathcal{F}$ is non-antipodal with respect to K.

Briefly, let me explain the terminology. We say that two metrics $d_1(\cdot,\cdot)$ and $d_2(\cdot,\cdot)$ on $\mathbb{R}^M$ are Lipshitz equivalent when there is a constant $C \geq 1$ such that $d_1(x,y)/C \leq d_2(x,y) \leq C d_1(x,y)$ for all $x, y \in \mathbb{R}^M$. Given any closed bounded set K in $\mathbb{R}^M$, whose affine hull is $\mathbb{R}^M$, and any $u \in S^{M-1}$, the unit sphere in $\mathbb{R}^M$, let $\{\mathcal{H}_u, \mathcal{H}_{-u}\}$ be the unique pair of distinct support hyperplanes of K perpendicular to u, see [27] p.14. Then the set of antipodal pairs of points of K is $K' := \left\{\{a, a'\} : a \in \mathcal{H}_u \cap \partial K,\ a' \in \mathcal{H}_{-u} \cap \partial K,\ u \in S^{M-1}\right\}$ where ∂K denotes the boundary of K. We say that an IFS $\mathcal{F}$ *is non-antipodal with respect to* K when each of its functions takes K into itself but maps no antipodal pair of points of K to an antipodal pair of points of K. We denote the latter condition by $\mathcal{F}(K') \cap K' = \emptyset$.

Part of the proof of Theorem 2 relies on the observation that if $\mathcal{K} \subset \mathbb{R}^M$ is a convex body (think of $\mathcal{K}$ as the convex hull of K in (3)) then we can define a metric $d_{\mathcal{K}}(\cdot,\cdot)$ on $\mathbb{R}^M$, Lipshitz equivalent to the Euclidean metric, by

$$d_{\mathcal{K}}(x,y) = \inf\left\{\frac{\|x-y\|}{\|l-m\|} : l, m \in \mathcal{K}, l - m = \alpha(x-y), \alpha \in \mathbb{R}\right\}$$

for all $x \neq y$, where $\|x-y\|$ denotes the Euclidean distance from x to y in $\mathbb{R}^M$. One shows that, if an affine map f_n is non-antipodal with respect to $\mathcal{K}$, then it is a contraction with respect to $d_{\mathcal{K}}$. In fact $d_{\mathcal{K}}$ is, up to a constant factor, a Minkowski

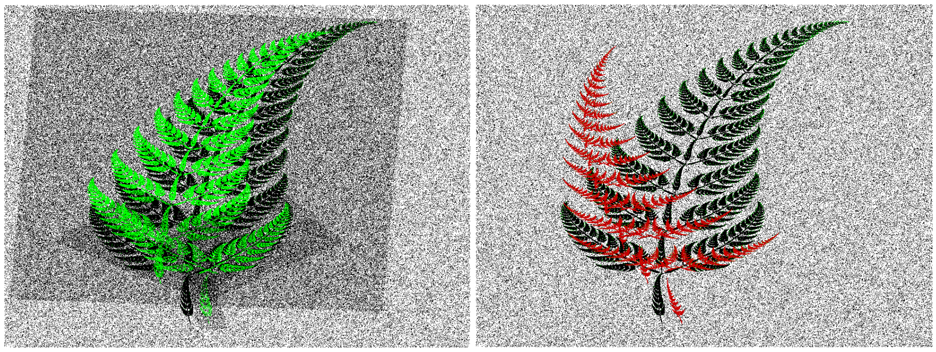


FIGURE 2. The left-hand panel shows a black fern image within a grimy window; overlayed upon it are four affine transformations of the window and the fern, with the transformed fern images shown in green. The goal has been to approximate the original (black) fern with affinely transformed (green) copies of itself. The rectangular window has been mapped non-antipodally upon itself. So, by Theorem 2, there exists a metric such that the associated affine IFS $\mathcal{F}$ is contractive. The original fern (black) and the attractor (red) of $\mathcal{F}$ are shown in the right-hand panel.

metric [30] associated with the symmetric convex body defined by the Minkowski difference $\mathcal{K} - \mathcal{K}$.

Theorem 2 provides a means for defining viable IFS codes (DNA) and is useful in the design of a two-dimensional affine IFS whose attractor approximates a given target set $T \subset \mathbb{R}^2$. Typical IFS software for this purpose exhibits a convex window $\mathcal{K}$, containing a picture of T, on a digital computer display. A set of affine maps is introduced, thereby defining an IFS $\mathcal{F}$. The maps are adjusted using interactive pictures of $\mathcal{F}(\mathcal{K})$ and $\mathcal{F}(T)$. If we ensure that $\mathcal{F}(\mathcal{K}) \subset \mathcal{K}$ and $\mathcal{F}(\mathcal{K}') \cap \mathcal{K}' = \emptyset$, then $\mathcal{F}$ possesses a unique attractor $A \subset \mathcal{K}$. An example is illustrated in Figure 2.

Now we are in a position to explain a stability relationship between IFS codes and attractors, and thus to exhibit a form of stability as in attribute (III). We can control the (Hausdorff) distance $h_{\mathcal{K}}$, which depends on $d_{\mathcal{K}}$, between A and T because it depends continuously on the distance between $\mathcal{F}(T)$ and T. Indeed, the collage theorem [2] states that

$$h_{\mathcal{K}}(A,T) \leq \frac{h_{\mathcal{K}}(\mathcal{F}(T),T)}{1-\lambda}$$

where $0 \leq \lambda < 1$ is a Lipshitz constant for $\mathcal{F} : \mathbb{H} \to \mathbb{H}$, for example the maximum of a set of contractivity factors of the f_ns with respect to $d_{\mathcal{K}}$. Notice that this relationship says nothing about the topological structure of an attractor: it comments only on its approximate shape.

The collage theorem expresses one kind of stability for the IFS creative system, as in attribute (III): small changes in the IFS code of a contractive IFS lead to small changes in the shape of the attractor. Indeed, this realization played a role in the development of fractal image compression [2]. In this development several things occurred. First, affine IFS theory adapted to a digital environment, illustrating a

component of attribute (V). Since contractive affine IFSs do not in general translate to contractive discrete operators, [28], new theory had to be developed. (See for example [16].) This illustrates stability of a second kind, as required in attribute (III): the underlying ideas are robust relative to structural changes in the creative system. Finally, a real application was the result, as required by attribute (V).

3. Projective and Bilinear IFSs

We will also use both projective and bilinear IFSs for creative applications that control the shape and topology of attractors, and transformations between attractors. Both are generalizations of two-dimensional affine IFSs. Both can be expressed with relatively succinct IFS codes, yet have more degrees of freedom than affines. Another such family of IFSs is provided by the Möbius transformations on $\mathbb{C}\cup\{\infty\}$. The availability of a rich selection of accessible examples is a valuable attribute for the survival of a mathematical idea.

In two dimensions, the functions of a *projective* IFS are represented in the form

$$f_n(x,y) = \left(\frac{a_n x + b_n y + c_n}{g_n x + h_n y + j_n}, \frac{d_n x + e_n y + k_n}{g_n x + h_n y + j_n}\right), \tag{3.1}$$

where the coefficients are real numbers. A similar result to Theorem 2 applies to such projective transformations restricted to a judiciously chosen convex body, with the associated Hilbert metric, [12] p.105, used in place of the generalized Minkowski metric mentioned above. Specifically, let $\mathcal{F}$ denote a projective IFS of the form $(\mathcal{K}^\circ; f_1, ..., f_N)$, where $\mathcal{K}^\circ$ is the interior of a convex body $\mathcal{K} \subset \mathbb{R}^2$ such that $\mathcal{F}(\mathcal{K}) \subset \mathcal{K}^\circ$. The associated Hilbert metric d_H is defined on $\mathcal{K}^\circ$ by

$$d_H(x,y) = \ln|\mathcal{R}(x,y;a,b)| \text{ for all } x,y \in \mathcal{K}^\circ \text{ with } x \neq y,$$

where $\mathcal{R}(x,y;a,b) = (|b-x|/|x-a|)/(|b-y|/|y-a|)$ denotes the cross ratio between x,y and the two intersection points a,b of the straight line through x,y with the boundary $\mathcal{K}$. You might like to verify that $\mathcal{F}$ is a contractive IFS with respect to d_H, using the fact that projective transformations preserve cross ratios. So projective IFSs can be used in applications in nearly the same way as affine systems.

To describe bilinear transformations, let $\mathcal{R} = [0,1]^2 \subset \mathbb{R}^2$ denote the unit square, with vertices $A=(0,0), B=(1,0), C=(1,1), D=(0,1)$. Let P,Q,R,S denote, in cyclic order, the successive vertices of a possibly degenerate quadrilateral. Then we uniquely define a *bilinear* function $\mathcal{B}:\mathcal{R}\to\mathcal{R}$ such that $\mathcal{B}(ABCD) = PQRS$ by

$$\mathcal{B}(x,y) = P + x(Q-P) + y(S-P) + xy(R+P-Q-S). \tag{3.2}$$

This transformation acts affinely on any straight line that is parallel to either the x-axis or the y-axis. For example, if $\mathcal{B}|_{AB}: AB \to PQ$ is the restriction of $\mathcal{B}$ to AB, and if $\mathcal{Q}:\mathbb{R}^2\to\mathbb{R}^2$ is the affine function defined by $\mathcal{Q}(x,y) = P + x(Q-P) + y(S-P)$, then $\mathcal{Q}|_{AB} = \mathcal{B}|_{AB}$. Because of this "affine on the boundary" property, bilinear functions are well suited to the construction of fractal homeomorphisms, as we will see. Sufficient conditions under which there exists a metric with respect to which a given bilinear transformation is contractive are given in [6]. A bilinear IFS has an attractor when its IFS code is close enough (in an appropriate metric) to the IFS code of an affine IFS that has an attractor.

An example of a geometrical configuration of quadrilaterals that gives rise to both a projective and a bilinear IFS is illustrated in Figure 3. In either case we

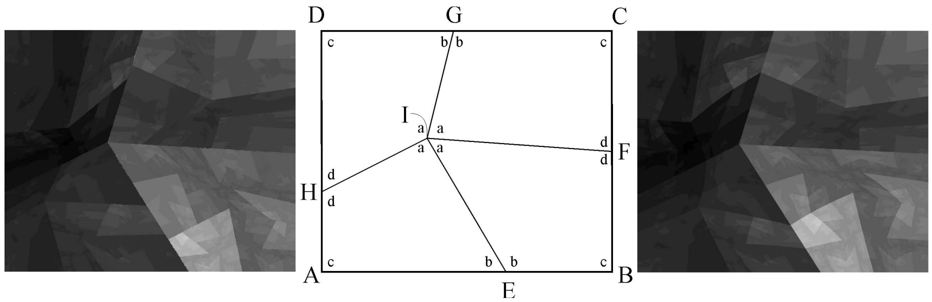


FIGURE 3. The four quadrilaterals $IEAH$, $IEBF$, $IGCF$, $IGDH$, define both an projective $\mathcal{F}$ and a bilinear IFS $\mathcal{G}$ both of which have a unique attractor, the filled rectangle with vertices at $ABCD$; but the address structures are different. In the left panel and right panels, the attractors of $\mathcal{F}$ and $\mathcal{G}$ respectively have been rendered, so that points with the same address are the same shade of grey. The address structure is independent of E, F, G, H, I, in the bilinear case but not in the projective case. (Hint: compare how lines meet the line IE.)

define the IFS to be $(\mathcal{R}; f_1, f_2, f_3, f_4)$ where

$$f_1(ABCD) = IEAH, f_2\,(ABCD) = IEBF, f_3\,(ABCD) = IGCF, f_3(ABCD) = IGDH,$$

where the first expression means $f_1\,(A) = I, f_1\,(B) = E, f_1\,(C) = A, f_1\,(D) = H$. With few constraints each IFS is contractive with respect to a metric that is Lipshitz equivalent to the Euclidean metric, with attractor equal to the filled rectangle $ABCD$. But there is an important difference: the bilinear family provides a family of homeomorphisms on $\mathcal{R}$, with applications to photographic art, attribute (V), while the projective family does not, as we will see.

4. The Chaos Game

How do we compute approximate attractors in a digital environment? Algorithms based on direct discretization of the expression $A = \lim_{k\to\infty} \mathcal{F}^{\circ k}(B)$ have high memory requirements and tend to be inaccurate, [28]. The availability of a simple algorithm that is fast and accurate, for the types of IFS that we discuss, has played an important role in the survival of the IFS creative system. The following algorithm, known as the *chaos game*, was described to a wide audience in *Byte Magazine* in 1988, and successfully dispersed IFS codes to the computer science community. It helped ensure that the IFS creative system would have attribute (V).

Define a *random orbit* $\{x_k\}_{k=0}^{\infty}$ of a point $x_0 \in \mathbb{X}$ under $\mathcal{F} = (\mathbb{X}; f_1, ..., f_N)$ by $x_k = f_{\sigma_k}(x_{k-1})$, where $\sigma_k \in \{1, 2, ..., N\}$ is chosen, independently of all other choices, by rolling an N-sided die. The random orbit is associated with $\sigma \in \Omega$. If the underlying space is two-dimensional and $\mathcal{F}$ is contractive, then it is probable that a picture of the attractor of $\mathcal{F}$, accurate to within viewing resolution, will be obtained by plotting $\{x_k\}_{k=100}^{10^7}$ on a digital display device.

Why does this Markov chain Monte Carlo algorithm work? The following theorem, implicit in [8], tells us how we can think of the attractor of a contractive IFS

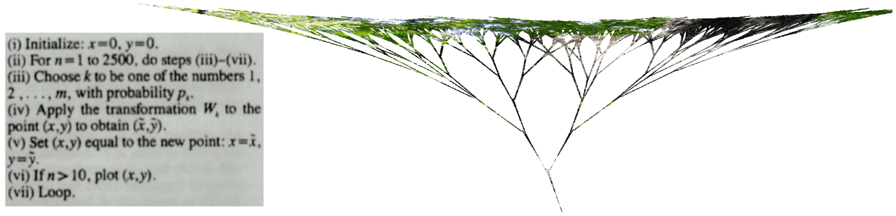

(i) Initialize: $x=0$, $y=0$.
(ii) For $n=1$ to 2500, do steps (iii)–(vii).
(iii) Choose k to be one of the numbers 1, 2, . . . , m, with probability p_k.
(iv) Apply the transformation W_k to the point (x,y) to obtain $(\tilde{x},\tilde{y})$.
(v) Set (x,y) equal to the new point: $x=\tilde{x}$, $y=\tilde{y}$.
(vi) If $n>10$, plot (x,y).
(vii) Loop.

FIGURE 4. The original pseudocode from *Byte Magazine* (Jan 1988) for implementing Chaos Game algorithm to obtain an image of the attractor of an IFS on $\mathbb{R}^2$. Notice the small number of iterations used! Nowdays, usually, I use 10^7 iterations, and discard the first thousand points. On the right is a sketch of a 2–variable tree obtained by a generalization of the chaos game.

FIGURE 5. From left to right this picture shows: the result of 9000 iterations of the chaos game algorithm applied to a projective IFS; the result of 10^7 iterations; a small picture of a flower; and a rendered close-up of the attractor. In the latter image the colours were obtained with the aid of a fractal transformation from the attractor to the small picture of the yellow flower.

as being the ω-limit set of almost any random orbit. A direct proof can be found in [31].

Theorem 3. *Let $\{x_k\}_{k=0}^{\infty}$ be a random orbit of a contractive IFS. With probability one*

$$\lim_{K\to\infty} \cup_{k=K}^{\infty} \{x_k\} = A,$$

where the limit is taken with respect to the Hausdorff metric.

Pictures, calculated using the chaos game algorithm, of the attractor of a projective IFS $(\mathcal{R}; f_1, f_2, f_3, f_4)$ are shown in the left-most two panels of Figure 5. You

can visualize an approximation to the stationary probability measure of the stochastic process, implicit in the chaos game, in the left-hand image. This measure depends on the strictly positive probabilities associated with the maps, but its support, the attractor, does not! Measure theory aspects of IFS are not considered in this article, but it is nice to see one way in which the topic arises.

5. Addresses, and transformations between attractors

In this section we deepen our understanding of attractors. We discover information about the relationship between IFS codes and the topology of attractors, and the relationships between different attractors. This information helps to classify our diverse plants, attribute (I), and leads into real applications in art and biology, attribute (V). Good book-keeping is the key.

The space $\Omega = \{1, 2, ..., N\}^{\infty}$ with the product topology plays a fundamental role in IFS theory and in this article. We write $\sigma = \sigma_1\sigma_2...$ to denote a typical element of Ω. We will use the notation $f_{\sigma_1\sigma_2...\sigma_k} := f_{\sigma_1} \circ f_{\sigma_2} \circ ... \circ f_{\sigma_k}$, $\sigma|k = \sigma_1\sigma_2...\sigma_k$, and $f_{\sigma_1\sigma_2...\sigma_k} = f_{\sigma|k}$ for any $\sigma \in \Omega$ and $k = 1, 2, ...$.

The following theorem suggests our plants can have intricate topological structures and suggests that symbolic dynamics are involved, thereby adding a lusher interpretation of attribute (I).

Theorem 4. [18] *Let $\mathcal{F} = (\mathbb{X}; f_1, ..., f_N)$ be a contractive IFS, with attractor A. Let $x \in \mathbb{X}$. A continuous surjection $\pi : \Omega \to A$, independent of x, is well-defined by $\pi(\sigma) = \lim_{k\to\infty} f_{\sigma|k}(x)$; the convergence is uniform for $(\sigma, x) \in \Omega \times B$, for any $B \in \mathbb{H}$.*

The set of addresses of a point $x \in A$ is defined to be the set $\pi^{-1}(x)$, and defines an equivalence relation $\sim$ on Ω. For example, the attractor of the IFS $(\mathbb{R}; f_1(x) = x/2, f_2(x) = x/2 + 1/2)$ is the closed interval $[0, 1]$. You may check that $\pi^{-1}(0) = \{\overline{1} := 1111...\}, \pi^{-1}(1) = \{\overline{2}\}$, $\pi^{-1}(1/2) = \{1\overline{2}, 2\overline{1}\}$, and $\pi^{-1}(1/3) = \{\overline{12}\}$. Some points of an attractor have one address while others have multiple distinct addresses. The topology on A is the identification topology on Ω induced by the continuous map $\pi : \Omega \to A$. In this paper we refer to the set of equivalence classess induced by $\sim$ on Ω as the *address structure* of the IFS. Figure 3 contrasts the address structures of a corresponding pair of bilinear and projective IFSs.

We can think of the topology of an attractor A as being that of Ω with all points in each equivalence class glued together; that is, A is homeomorphic to $\Omega/\sim$. Simple examples demonstrate that the address structure can change in complicated ways when a single parameter is varied: the topologies of attractors, our plants, in contrast to their shapes, do not in general depend continuously on their IFS codes. By restricting to appropriate families of projective or bilinear IFSs, with known address structures, control of the topology of attractors becomes feasible.

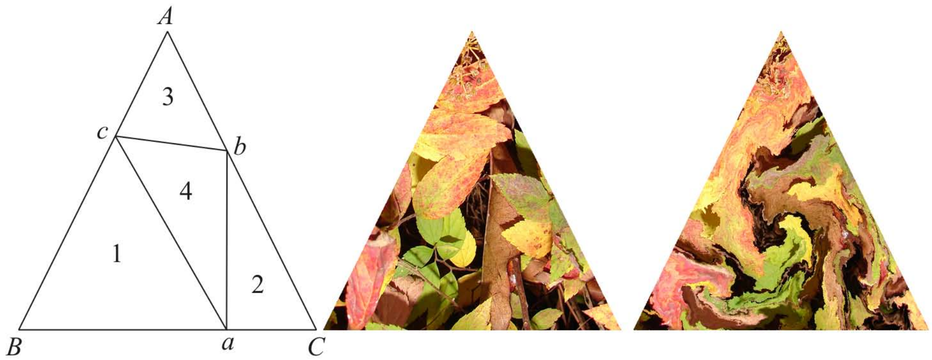
A point on an attractor may have multiple addresses. We select the "top" address to provide a unique assignment; the *top address* is the one closest to $\overline{1} = 1111...$ in lexographic ordering. Each element of the address structure of an IFS is represented by a unique point in Ω. This choice is serendipitous, because the resulting set of addresses, called the tops space, is shift invariant, and so yields a link between our plants, symbolic dynamics, and information theory, attribute (IV), see [5] and references therein.

FIGURE 6. Three renderings of a close-up of the attractor in Figure 5 computed using a coupled version of the chaos game and a fractal transformation. On the left the computation has been stopped early, yielding a misty effect. Different aspects of the fractal transformation are revealed by applying it to different pictures.

We define a natural map from an attractor $A_{\mathcal{F}}$ of an IFS $\mathcal{F}$ to the attractor $A_{\mathcal{G}}$ of an IFS $\mathcal{G}$, each with the same number of maps, by assigning to each point of $A_{\mathcal{F}}$ the point of $A_{\mathcal{G}}$ whose set of addresses includes the top address of the point in $A_{\mathcal{F}}$. This provides a map $T_{\mathcal{FG}} : A_{\mathcal{F}} \to A_{\mathcal{G}}$ called a *fractal transformation.* When the address structures of $A_{\mathcal{F}}$ and $A_{\mathcal{G}}$ are the same, this map is a homeomorphism. Since fractal transformations can be readily computed by means of a coupled version of the chaos game, applications to art and geometric modelling become feasible and the IFS creative system tests new forms and environments, attribute (V).

Let $\mathcal{R} \subset \mathbb{R}^2$ denote a filled unit square. Let $p : \mathcal{R} \to \mathcal{C}$ be a *picture (function)*; that is, p is a mapping from $\mathcal{R}$ into a color space $\mathcal{C}$. A *color space* is a set of points each of which is associated with a unique color. In computer graphics a typical color space is $\mathcal{C} = \{0, 1, ..., 255\}^3$, where the coordinates of a point represent digital values of red, green and blue. The graph of a picture function may be represented by a colorful picture supported on $\mathcal{R}$. Next time you see a picture hanging on a wall, imagine that it is instead, an abstraction, a graph of a picture function. More generally, we allow the domain of a picture function to be an arbitrary subset of $\mathbb{R}^2$.

For example, in the right-hand image in Figure 5 we have rendered the graph of $\widetilde{p} : A_{\mathcal{F}} \to \mathcal{C}$ obtained by choosing p to correspond to the picture of the yellow flower, $\mathcal{G}$ to be an affine IFS such that $\{g_n(\mathcal{R})\}_{n=1}^4$ is a set of rectangular tiles with $\cup g_n(\mathcal{R}) = \mathcal{R}$, and $\mathcal{F}$ to be the projective IFS whose attractor is illustrated in black. In Figure 6 we show other renderings of a portion of the attractor, obtained by changing the picture p and, in the left-hand image, by stopping the chaos game algorithm early–the pictures are all computed by using a coupled variant of the chaos game. Since $\widetilde{p} = p \circ T_{\mathcal{FG}}$ one can infer something about the nature of fractal transformations looking at such pictures. By panning the source picture p it is possible to make fascinating video sequences of images. You can see some yourself with the aid of *SFVideoShop* [32]. In the present example you would quickly infer that $T_{\mathcal{FG}}$ is not continuous but that it is not far from being so: it may be continuous except across a countable set of arcs.



(i) (ii) (iii)

FIGURE 7. (i) The adjustable points a, b, c used to define a family of affine iterated function systems $\mathcal{F} = (\mathbb{R}^2; f_1, f_2, f_3, f_4)$ with constant address structure; (ii) a picture supported on the attractor of an IFS belonging to the family; (iii) the same picture, transformed under a fractal homeomorphism of the form $\mathcal{T}_{\mathcal{FG}} = \pi_{\mathcal{G}} \circ \tau_{\mathcal{F}}$ where $\mathcal{F}$ and $\mathcal{G}$ are two IFSs belonging to the family.

6. The art of fractal homeomorphism

Affine and bilinear iterated function systems can be used to provide a wide variety of parameterized families of homeomorphisms on two-dimensional regions with polygonal boundaries such as triangles and quadrilaterals. We use them to illustrate the application of the IFS creative system to a new artform. In effect this application is itself a new creative system, for artists. This provides an illustration of attribute (VI): creative systems beget creative systems.

For example, let A, B, and C denote three non-colinear points in $\mathbb{R}^2$. Let c denote a point on the line segment AB, let a denote a point on the line segment BC, and let b denote a point on the line segment CA, such that $\{a, b, c\} \cap \{A, B, C\} = \varnothing$; see Figure 7(i). Let $\mathcal{F} = \left(\mathbb{R}^2; f_1, f_2, f_3, f_4\right)$ be the unique affine IFS such that

$$f_1(ABC) = caB,\ f_2(ABC) = Cab,\ f_3(ABC) = cAb, \text{ and } f_4(ABC) = cab,$$

where we mean for example that f_1 maps A to c, B to a, and C to B, see Figure 7(i). For reference, let us write $\mathcal{F} = \mathcal{F}_{\alpha,\beta,\gamma}$ where $\alpha = |Bc|/|AB|, \beta = |Ca|/|BC|$, and $\gamma = |Ab|/|CA|$. The attractor of $\mathcal{F}_{\alpha,\beta,\gamma}$ is the filled triangle $\mathcal{T}$ with vertices at A, B, and C. Then $\mathcal{F}_{\alpha,\beta,\gamma}$ is contractive IFS, for each $(\alpha, \beta, \gamma) \in (0, 1)^3$, with respect to a metric that is Lipshitz equivalent to the Euclidean metric, by Theorem 2. Its address structure $\mathcal{C}_{\alpha,\beta,\gamma}$ is independent of α, β, γ, see [5] section 8.1.

Figure 7(ii) illustrates a picture $p : \mathcal{T} \to \mathcal{C}$; it depicts fallen autumn leaves. Figure 7(iii) illustrates the picture $\widetilde{p} = p \circ T_{\mathcal{FG}}$, namely the result of applying the homeomorphism $T_{\mathcal{FG}}$ to the picture p, where $\mathcal{F} = \mathcal{F}_{0.45,0.45,0.45}$ and $\mathcal{G} = \mathcal{F}_{0.55,0.55,0.55}$. The transformation in this example is area-preserving because corresponding tiles have equal areas.

A similar result applies to families of bilinear IFSs. For example Figure 3 defines a family of bilinear IFSs, $\mathcal{F}_v$, parameterized by the vector of points $v = (E, F, G, H, I)$. This family has constant address structure for all values of v for

FIGURE 8. Example of a fractal homeomorphism generated by two IFSs of four bilinear transformations. The attractor of each transformation is the unit square.

FIGURE 9. Before (left) and after a fractal homeomorphism.

which $\mathcal{F}_v$ is contractive and can thus be used to provide a family of homeomorphisms $\mathcal{T}_{v,w} : \mathcal{R} \to \mathcal{R}$. An illustration of the action of $\mathcal{T}_{v,w}$ on a picture of Australian heather is given in Figure 8. In this case the parameters v and w both correspond to affine IFSs. What is remarkable in this case, and many like it, is that the transformed picture looks so realistic. Can you tell which is the original?

Figure 9 illustrates a homeomorphic fractal transformation generated by a pair of bilinear IFSs on $\mathcal{R}$. In this case $N = 12$. The original image is a digital photograph of a lemon tree and wallflowers in my garden in Canberra. The final image was printed out on thick acid-free rag paper by a professional printing company, using vivid pigment inks, at a width of approximately 5ft and a height of 3ft 6ins. It represents a fusion of the colors of nature and mathematics; it provokes wonder in me, a sense of the pristine and inviolate, a yearning to look and look ever closer, see Figure 10.

I have used such extraordinary transformations to generate works for three successful (most of the pictures are sold) art shows, in Canberra (Australia, April 2008), in Bellingham (Washington State, April 2008), and in Gainesville (Florida, March 2009).

FIGURE 10. Detail from Figure 9 showing not only the vibrant colors of nature, but also the wide range of stretching and squeezing achieved in this relatively simple fractal transformation.

7. Superfractals

In this section we illustrate attribute (VI). We show how IFS theory begets a new creative system via a higher level of abstraction. The new framework is suitable for mathematical modelling of the geometry of a multitude of naturally occurring, readily observable structures. It also has applications to the visual arts.

The new system has some remarkable properties. Its attractor is a set of interrelated sets that can be sampled by a variant of the chaos game algorithm, as illustrated in Figure 12. This algorithm is born fully formed and is the key to applications. The geometry and topology of the interrelated sets can be controlled when appropriate generalized IFSs are used. In particular, through the concepts of V-variability [7] and superfractals, we are able to form a practical bridge between deterministic fractals (such as some of the IFS attractors in previous sections) and random fractal objects (such as statistically self similar curves that represent Brownian motion).

7.1. $V-$**variability.** Here is a biological way to think of "$V-$variability". Imagine a tree that grows with this property. If you were to break off all of the branches of any one generation and classify them, you would find that they were of at most V

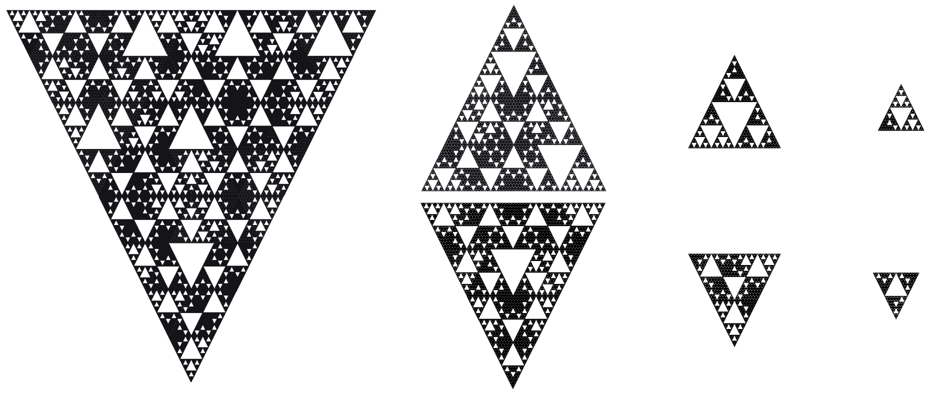


FIGURE 11. An element of a 2-variable superfractal is shown on the left. If you look closely at it, you will see that it is made of exactly two distinct(up to translation, reflection and rotation) sub-objects of half the linear dimension. (The two objects immediately to the right of the first one.) And, if you look even closer, you will see that it is made of two sub-objects of one-quarter the linear dimension, and so on.

different types. By "generation" I mean that you are able to think of the tree as having older and younger branches, that is, some that started to grow during year one, sub-branches that began during year two and so on. The tree is very old. By "type" I mean something like "belongs to a particular conjugacy class". The type may change from generation to generation, but the number V is fixed and as small as possible. Then I will call the imagined tree "V–variable". Figure 4 includes an illustration of a 2–variable tree, where the younger branches start higher up the tree. Again, consider a population of annual plants belonging to a species that admits S distinct possible genotypes. If the number of distinct genotypes in each generation is bounded above by V then (in circumstances where V is significantly smaller than S) I would call this population "V–variable". But the mathematical definition relates to a property of attractors of certain IFSs. Figure 11 illustrates a 2–variable fractal subset of the Euclidean plane: it is a union of two tiles of half its size: it is also a union of at most two tiles of a quarter its size, and so on.

Let $\mathcal{F} = (\mathbb{X}; f_1, ..., f_N)$ be an IFS of functions f_n that are contractive with respect to the metric d on $\mathbb{X}$. If $\mathcal{G} = (\mathbb{X}; f_{\omega_1}, .., f_{\omega_l})$ for some choice of indices $1 \leq \omega_1 < \omega_2 < ... < \omega_l \leq N$, then we say that $\mathcal{G}$ is a *subIFS* of $\mathcal{F}$.

Given an IFS $\mathcal{G} = (\mathbb{X}; g_1, ..., g_M)$ and a sequence of indices $\rho = \rho_1\rho_2...\rho_M$, where each ρ_m belongs to $\{1, 2, ..., V\}$ we can construct a mapping $\mathcal{G}^{(\rho)} : \mathbb{H}^V \to \mathbb{H}$ by defining

$$\mathcal{G}^{(\rho)}(\underline{B}) = \cup_m g_m \left(B_{\rho_m}\right), \text{ for all } \underline{B} = (B_1, B_2, ..., B_V) \in \mathbb{H}^V.$$

In a similar manner, given a set of subIFSs $\{\mathcal{G}_1, ..., \mathcal{G}_L\}$ of $\mathcal{F}$, each consisting of M functions, we can construct mappings from $\mathbb{H}^V$ to itself. Let $\sigma = \sigma_1\sigma_2...\sigma_V \in \{1, 2, ..., L\}^V$, let $\boldsymbol{\rho}$ be a $V \times M$ matrix whose entries belong to $\{1, 2, ..., V\}$, and here let ρ_v denote the v^{th} row of $\boldsymbol{\rho}$. Then we define a mapping $\mathcal{G}^{(\boldsymbol{\rho},\sigma)} : \mathbb{H}^V \to \mathbb{H}^V$

by

$$\mathcal{G}^{(\boldsymbol{\rho},\sigma)}(\underline{B})=(\mathcal{G}_{\sigma_1}^{(\boldsymbol{\rho}_1)}(\underline{B}),\mathcal{G}_{\sigma_2}^{(\boldsymbol{\rho}_2)}(\underline{B}),...,\mathcal{G}_{\sigma_V}^{(\boldsymbol{\rho}_V)}(\underline{B})).$$

We denote the sequence of all such mappings by $\{\mathcal{H}_j : j \in J\}$ where J is the set of all indices $(\boldsymbol{\rho},\sigma)$, in some order. We call $\mathfrak{G}^{(V)} = \left(\mathbb{H}^V; \{\mathcal{H}_j : j \in J\}\right)$ the *V–variable superIFS* associated with the set of subIFSs $\{\mathcal{G}_l\}_{l=1}^{L}$ of $\mathcal{F}$.

We write B_v to denote the v^{th} component of $B \in \mathbb{H}^V$. If the space $\mathbb{H}^V$ is equipped with the metric $D(B,C) := \max_v \{h(B_v, C_v)\}$ where h is the Hausdorff metric on $\mathbb{H}$, then $\left(\mathbb{H}^V, D\right)$ is a complete metric space. The following theorem summarizes basic information about $\mathfrak{G}^{(V)}$. More information is presented in [4] and [7].

Theorem 5. [7] *Let $\mathfrak{G}^{(V)}$ denote the $V-$variable superIFS $\left(\mathbb{H}^V; \{\mathcal{H}_j : j \in J\}\right)$.*

(i) If the underlying IFS $\mathcal{F}$ is contractive then the IFS $\mathfrak{G}^{(V)}$is contractive.

(ii) The unique attractor $\mathbb{A}^{(V)} \in \mathbb{H}\left(\mathbb{H}^V\right)$ of $\mathfrak{G}^{(V)}$ consists of a set of V–tuples of compact subsets of $\mathbb{X}$, and $A^{(V)} := \{B_v : B \in \mathbb{A}^{(V)}, v = 1, 2, ..., V\} = \{B_v : B \in \mathbb{A}^{(V)}\}$ for all $v = 1, 2, ..., V$. (Symmetry of the superIFS with respect to the V coordinates causes this.) Each element of $A^{(V)}$ is a union of transformations, belonging to $\mathcal{F}$, of at most V other elements of $A^{(V)}$.

(iii) If $\{\mathbb{A}_k\}_{k=0}^{\infty}$ denotes a random orbit of $\mathbb{A}_0 \in \mathbb{H}^V$ under $\mathfrak{G}^{(V)}$ and $A_k \in \mathbb{H}$ denotes the first component of $\mathbb{A}_k$, then (with probability one) $\lim_{K\to\infty} \cup_{k=K}^{\infty} \{A_k\} = A^{(V)}$ where the limit is taken with respect to the Hausdorff metric on $\mathbb{H}(\mathbb{H}(\mathbb{X}))$.

Statement (i) implies that $\mathfrak{G}^{(V)}$ possesses a unique attractor $\mathbb{A}^{(V)}$ and that we can describe it in terms of the chaos game. This technique is straightfoward to apply, since we need only to select independently at each step, the indices $\boldsymbol{\rho}$ and σ; the functions themselves are readily built up from those of the underlying IFS $\mathcal{F}$.

Statement (ii) tells us that it is useful to focus on the set $A^{(V)}$ of first components of elements of $\mathbb{A}^{(V)}$. It also implies that, given any $A \in A^{(V)}$ and any positive integer K, there exist $A_1, A_2, ..., A_V \in A^{(V)}$ such that $A = \cup_{l\in\{1,2,...,V\}} \cup_{\sigma\in\Omega} f_{\sigma|K}(A_l)$. That is, at any depth K, A is a union of contractions applied to V sets, all belonging to $A^{(V)}$, at most V of which are distinct. In view this property, the elements of $A^{(V)}$ are called *V–variable fractal sets* and we refer to $A^{(V)}$ itself as a *superfractal*.

Statement (iii) tells us that we can use random orbits $\{\mathbb{A}_k\}_{k=0}^{\infty}$ of $\mathbb{A}_0 \in \mathbb{H}^V$ under $\mathfrak{G}^{(V)}$ to sample the superfractal $A^{(V)}$.

The idea of address structures, tops and fractal transformations can be extended to the individual sets that comprise $A^{(V)}$, see [4]. We are thus able to render colorful images of sequences of elements of $A^{(V)}$ generated by a more elaborate chaos game involving, at each step a $(V + 1)$–tuple of sets, one of which is used to define the picture whose colors are used to render the other sets.

Figure 12 illustrates elements taken from such an orbit. In this case a 2–variable superIFS is used: it consists of two subIFS of a projective IFS $\mathcal{F}$, consisting of five functions, detailed in [4]. The fern-like structure of all elements of the corresponding superfractal is ensured by a generalized version of the collage theorem.

7.2. **The problem solved by superfractals.** The diverse forms that illustrate a successful idea, the plants of a creative system, may change as time goes forward. The ideas of an earlier era of geometry, popular in applications included cissoids, strophoids, nephroids, and astroids: more recently you would hear about manifolds, Ricci curvature, and vector bundles; today you are just as likely to hear about

FIGURE 12. Elements of a 2–variable superfractal, rendered as in Figure 5.

fractals. Why? As technology advances, some applications become extinct, and new ones emerge.

In *The Fractal Geometry of Nature*, [24], Mandelbrot argues that random fractals provide geometrical models for naturally occurring shapes and forms, such as coastlines, clouds, lungs, trees, and Brownian motion. A *random fractal* is a statistically self-similar object with non-integer Hausdorff dimension. Although there are mathematical theories for families of random fractals, see for example [25], they are generally cumbersome to use in geometric modelling applications.

For example, consider the problem of modelling real ferns: ferns look different at different levels of magnification, and the locations of the fronds are not according to some strictly deterministic pattern, as in a geometric series, but rather they have elements of randomness. It seems that a top-down hierarchical description, starting at the coarsest level, and working down to finer scales, is needed to provide specific geometrical information about structure at all levels of magnification. This presents a problem: clearly it is time consuming and expensive in terms of the amount of data needed, to describe even a single sample from some statistical ensemble of such objects.

Superfractals solve this problem by restricting the type of randomness to be V-variable. This approach enables the generalized chaos game algorithm, described above, to work, yielding sequences of samples from a probability distribution on V-variable sets belonging to a superfractal. In turn, this means that we can approximate fully random fractals because, in the limit as V tend to infinity, V-variable fractals become random fractals in the sense of [25], see for example [7], Theorem 51.

Thus, we are able to compute arbitrarily accurate sequences of samples of random fractals. Furthermore we have modelling tools, obtained by generalizing those that belong to IFS theory, such as collage theorem and fractal transformations, which extend in natural ways to the V-variable setting. In some cases the Hausdorff dimension of these objects can also be specified as part of the model. This provides an approach to modelling many naturally occurring structures that is both mathematically satisfying and computationally workable. In particular, we see how the IFS creative system begat a new, even more powerful system, with diverse potential applications. This completes my argument that iterated function systems comprise a creative system.

8. Further reading

I would have liked to tell you much more about IFS theory. But this is not a review article, even of some of my own work. It does not touch the full range of the subject, let alone the mathematics of fractal geometry as a whole. The contents were chosen primarily to illustrate the idea of a mathematical creative system.

To survey mathematical fractal geometry, I mention the series of four conference proceedings [36], [37], [38], and [39], carefully edited by Christoph Bandt, Martina Zähle and others. The books by Falconer, for example [14] and [15], are good textbooks for core material. A recent development has been the discovery of how to construct harmonic functions and a calculus on certain fractal sets, see [20]. This was reported in the *Notices* [33]. A light introduction is [34]. Fractals and number theory is an important area; see for example [22], [11], and [23]. The topic of *noncommutative fractal geometry* is another fascinating new area [21]. Fractal geometry is rich with creative possibilities.

9. Conclusion

In this essay I have illustrated the notion that mathematical ideas that survive are creative systems in their own right, with attributes that parallel some of natural evolution.

Creative systems define, via their DNA, diverse forms and structures. There are three concepts here: seeds, plants, and diversity. Individual plants are products of the system, representatives of its current state and utility. The system itself may remain constant but the plants evolve, adapting to new generations of minds. The IFS creative system lives in my mind. But mainly I watch its plants: ones that preoccupy me now are not the same as the ones that I looked at years ago; the crucial element is the creative system, not the fractal fern.

Plants provide the first wave of conquest of new environments; an adapted version of the underlying new idea may follow later. The diversity of plants suggests a multitude of possibilities. Their seeds get into the minds of engineers and scientists. Later, the underlying idea, the creative system itself, may take hold.

I think of a good mathematical mind, a strong mathematics department, and a successful conference, as each being like Darwin's bank, a rich ecosystem, a fertile environment where ideas interact and diverse species of plants are in evidence. Some of these plants may be highly visible because they can be represented using computer graphics, while others are more hidden: you may only see them in colloquia, a few glittering words that capture and describe something wonderful, jump from brain to brain and there take root. (I think of the first time I heard about the Propp-Wilson algorithm.)

A good mathematical idea is a creative system, a source of new ideas, as rich in their own right as the original. The idea that survives is one that takes root in the minds of others: it does so because it is accessible and empowering. Such an idea is likely to lead to applications, but this applicability is more a symptom that the idea is a creative system, rather than being causative. A good idea allows, invites, surprises, simplifies, and shares itself without ever becoming smaller; as generous, mysterious, and bountiful as nature itself.

ACKNOWLEDGEMENTS. I acknowledge and thank David C. Wilson, Örjan Stenflo and Andrew Vince for helpful comments. I thank Louisa Barnsley for the illustrations.

References

[1] Ross Atkins, M. F. Barnsley, David C. Wilson, Andrew Vince, A characterization of point-fibred affine iterated function systems, submitted for publication, (2009) *arXiv:0908.1416v1*.

[2] M. F. Barnsley, Fractal image compression, *Notices Am. Math. Soc.* **43** (1996) 657-662.

[3] ———, Theory and application of fractal tops. In *Fractals in Engineering:New Trends in Theory and Applications,* J. Lévy-Véhel; E. Lutton (eds.) pp.3-20. London, Springer-Verlag, 2005.

[4] ———, *Superfractals*, Cambridge University Press, Cambridge, 2006.

[5] ———, Transformations between self-referential sets, *American Mathematical Monthly,* **116** (2009) 291-304.

[6] ———, Transformations between fractals, *Progress in Probability,* **61** (2009) 227-250.

[7] M. F. Barnsley, J. Hutchinson, Ö. Stenflo, V-variable fractals: fractals with partial self similarity, *Advances in Mathematics*, **218 (**2008) 2051-2088.

[8] M. F. Barnsley and S. G. Demko, Iterated function systems and the global construction of fractals, *Proc. Roy. Soc. London Ser. A* **399** (1985) 243–275.

[9] M. A. Berger and Y. Wang, Bounded semigroups of matrices, *Linear Algebra and Appl.* **166** (1992), 21-27.

[10] I. Daubechies and J.C. Lagarias, Sets of matrices all infinite products of which converge, *Linear Algebra and Appl.* **162** (1992), 227-263.

[11] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics,* V. Berthé, S. Ferenczi, C. Mauduit (eds.), *Lecture Notes in Mathematics* **1794** (2002) Springer-Verlag, Berlin.

[12] Herbert Buseman, *The Geometry of Geodesics,* Academic Press, New York, 1955.

[13] L. Janos, A converse of Banach's contraction theorem, *Proc. Amer. Math. Soc.* **18** (1967) 287–289.

[14] Kenneth Falconer, *Fractal Geometry: Mathematical Foundations and Applications* (Second Edition), John Wiley & Sons, Chichester, 2003.

[15] Kenneth Falconer, *Techniques in Fractal Geometry,* Cambridge University Press, Cambridge, 1997.

[16] Yuval Fisher, *Fractal Image Compression: Theory and Application,* Springer Verlag, New York, 1995.

[17] M. W. Hirsh, S. Smale, *Differential Equations and Linear Algebra*, Academic Press, Inc., Harcourt Brace Jovanonvich, San Diego, 1974.

[18] J. E. Hutchinson, Fractals and self-similarity*, Indiana Univ. Math. J.* **30** (1981) 713–747.

[19] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems. With a Supplementary Chapter by Katok and Leonardo Mendoza,* Cambridge University Press, Cambridge, 1995.

[20] J. Kigami, *Harmonic Calculus on p.c.f. Self-similar Sets,* Trans. Amer. Math. Soc. **335** (1993) 721-755.

[21] Michel L. Lapidus, Towards a noncommutative fractal geometry? *Contemporary Mathematics,* **208** (1997) 211-252.

[22] Michel L. Lapidus and Machiel van Frankenhuysen, *Fractal Geometry and Number Theory,* Birkhäuser, Boston, 1999.

[23] Wolfgang Müller, Jörg M. Thuswaldner, and Robert Tichy, Fractal Properties of Number Systems, *Period. Math. Hungar* **42** (2001) 51-68.

[24] Benoit Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, 1983.

[25] R. Daniel Mauldin and S. C. Williams, Random recursive constructions: asymptotic geometrical and topological properties. *Trans. Amer. Math. Soc.* **295** (1986), 325-346.

[26] B. Mendelson, *Introduction to Topology* (British edition), Blackie & Son, London, 1963.

[27] Maria Moszyńska, *Selected Topics in Convex Geometry,* Birkhäuser, Boston, 2006.

[28] Mario Peruggia, *Discrete Iterated Function Systems,* A.K. Peters, Wellesley, MA, 1993.

[29] A. Rényi, Representations of real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar.* **8** (1957), 477-493.

[30] R. Tyrrel Rockafellar, *Convex Analysis,* Princeton University Press, Princeton, 1970.

[31] Robert Scealy, *V-variable Fractals and Interpolation*, Ph.D Thesis, Australian National University, 2008.

[32] *SFVideoShop,* Windows software available free from www.superfractals.com. Written by M. F. Barnsley and R. Xie.

[33] Robert S. Strichartz, Analysis on fractals, *Notices Amer. Math. Soc.,* **46** (1999) 1199-1208.

[34] Robert S. Strichartz, *Differential Equations on Fractals*, Princeton University Press, 2006.

[35] Stephen Wolfram, *A New Kind of Science*, Wolfram Media, Inc., Champagne, 2002.

[36] Christoph Bandt, Siegfried Graf, Martina Zähle, (eds.) *Fractal Geometry and Stochastics, (Progress in Probability* **37**) Birkhäuser Verlag, Basel, 1995.

[37] Christoph Bandt, Siegfried Graf, Martina Zähle, (eds.) *Fractal Geometry and Stochastics II, (Progress in Probability* **46**) Birkhäuser Verlag, Basel, 2000.

[38] Christoph Bandt, Umberto Mosco, Martina Zähle, (eds.) *Fractal Geometry and Stochastics III, (Progress in Probability* **57**) Birkhäuser Verlag, Basel, 2004.

[39] Christoph Bandt, Peter Mörters, Martina Zähle, (eds.) *Fractal Geometry and Stochastics IV, (Progress in Probability* **61**) Birkhäuser Verlag, Basel, 2009.

Mathematical Sciences Institute, Australian National University

E-mail address: `mbarnsley@aol.com michael.barnsley@anu.edu.au`

URL: `http://www.superfractals.com`

A CHARACTERIZATION OF HYPERBOLIC AFFINE ITERATED FUNCTION SYSTEMS

ROSS ATKINS, MICHAEL F. BARNSLEY, ANDREW VINCE, AND DAVID C. WILSON

Abstract. The two main theorems of this paper provide a characterization of hyperbolic affine iterated function systems defined on $\mathbb{R}^m$. Atsushi Kameyama (*Distances on Topological Self-Similar Sets,* Proceedings of Symposia in Pure Mathematics, Volume **72.1,** 2004) asked the following fundamental question: given a topological self-similar set, does there exist an associated system of contraction mappings? Our theorems imply an affirmative answer to Kameyama's question for self-similar sets derived from affine transformations on $\mathbb{R}^m$.

1. Introduction

The goal of this paper is to prove and explain two theorems that classify hyperbolic affine iterated function systems defined on $\mathbb{R}^m$. One motivation was the question: when are the functions of an affine iterated function systems (IFS) on $\mathbb{R}^m$ contractions with respect to a metric equivalent to the usual euclidean metric?

Theorem 1.1 (Classification for Affine Hyperbolic IFSs)**.** *If $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is an affine iterated function system, then the following statements are equivalent.*

1. *$\mathcal{F}$ is hyperbolic.*
2. *$\mathcal{F}$ is point-fibred.*
3. *$\mathcal{F}$ has an attractor.*
4. *$\mathcal{F}$ is a topological contraction with respect to some convex body $K \subset \mathbb{R}^m$.*

Date: June 30, 2009.
2000 *Mathematics Subject Classification.* Primary 54H25, 26A18, 28A80.
Key words and phrases. iterated function systems, affine mappings, hyperbolic IFS, contraction mapping.

(5) *$\mathcal{F}$ is non-antipodal with respect to some convex body $K \subset \mathbb{R}^m$.*

Statement (1) is a metric condition on an affine IFS, statements (2) and (3) are in terms of convergence, and statements (4) and (5) are in terms of concepts from convex geometry. The terms contractive, hyperbolic, point-fibred, attractor, topological contraction, and non-antipodal are defined in Definitions 2.2, 2.3, 2.5, 2.7, 5.8, 6.5, respectively. This theorem draws together some of the main concepts in the theory of iterated function systems. Banach's classical Contraction Mapping Theorem states that a contraction f on a complete metric space has a fixed point x_0 and that $x_0 = \lim_{k\to\infty} f^{\circ k}(x)$, independent of x, where $\circ k$ denotes the k^{th} iteration. The notion of hyperbolic generalizes to the case of an IFS the contraction property, namely an IFS is hyperbolic if there is a metric on $\mathbb{R}^m$, equivalent to the usual one, such that each function in the IFS is a contraction. The notion of point-fibred, introduced by Kieninger [9], is the natural generalization of the limit condition above to the case of an IFS. While traditional discussions of fractal geometry focus on the existence of an attractor for a hyperbolic IFS, Theorem 1.1 establishes that the more geometrical (and non-metric) assumptions - topologically contractive and non-antipodal - can also be used to guarantee the existence of an attractor. Basically a function $f : \mathbb{R}^m \to \mathbb{R}^m$ is non-antipodal if certain pairs of points (antipodal points) on the boundary of K are not mapped by f to another pair of antipodal points.

Since the implication (1) $\Rightarrow$ (2) is the Contraction Mapping Theorem when the IFS contains only one affine mapping, Theorem 1.1 contains an affine IFS version of the converse to the Contraction Mapping Theorem. Thus, our theorem provides a generalization of results proved by L. Janos [7] and S. Leader [11]. Such a converse statement in the IFS setting has remained unclear until now.

Although not every affine IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is hyperbolic on all of $\mathbb{R}^m$, the second main result states that if $\mathcal{F}$ has a coding map (Definition 2.4), then $\mathcal{F}$ is always hyperbolic on some affine subspace of $\mathbb{R}^m$.

Theorem 1.2. *If $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is an affine IFS with a coding map $\pi : \Omega \to \mathbb{R}^m$, then $\mathcal{F}$ is hyperbolic on the affine hull of*

$\pi(\Omega)$. In particular, if $\pi(\Omega)$ contains a non-empty open subset of $\mathbb{R}^m$, then $\mathcal{F}$ is hyperbolic on $\mathbb{R}^m$.

Although he used slightly different terminology, Kameyama [8] posed the following *FUNDAMENTAL QUESTION*: *Is an affine IFS with a coding map $\pi : \Omega \to \mathbb{R}^m$ hyperbolic when restricted to $\pi(\Omega)$?* An affirmative answer to this question follows immediately from Theorem 1.2.

Our original motivation, however, was not Kameyama's question, but rather a desire to approximate a compact subset $T \subset \mathbb{R}^m$ as the attractor A of an iterated function system $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ...f_N)$, where each $f_n : \mathbb{R}^m \to \mathbb{R}^m$ is affine. This task is usually done using the "collage theorem" [1], [2] by choosing an IFS $\mathcal{F}$ so that the Hausdorff distance $d_{\mathbb{H}}(T, \mathcal{F}(T))$ is small. If the IFS $\mathcal{F}$ is hyperbolic, then we can guarantee it has an attractor A such that $d_{\mathbb{H}}(T, A)$ is comparably small. But then the question arises: how does one know if $\mathcal{F}$ is hyperbolic?

The paper is organized as follows. Section 2 contains notation, terminology, and definitions that will be used throughout the paper. Section 3 contains examples and remarks relating iterated function systems and their attractors to Theorems 1.1 and 1.2. In Example 3.1 we show that an affine IFS can be point-fibred, but not contractive under the usual metric on $\mathbb{R}^m$. Thus, some kind of remetrization is required for the system to be contractive. In Example 3.2 we show that an affine IFS can contain two linear maps each with real eigenvalues all with magnitudes less than 1, but still may not be point-fibered. Thus, Theorem 1.1 cannot be phrased only in terms of eigenvalues and eigenvectors of the individual functions in the IFS. Indeed, in Example 3.3 we explain how, given any integer $M > 0$, there exists a linear IFS $\left(\mathbb{R}^2; L_1, L_2\right)$ such that each operator of the form $L_{\sigma_1} L_{\sigma_2} ... L_{\sigma_k}$, with $\sigma_j \in \{1, 2\}$ for $j = 1, 2, ..., k$, and $k \leq M$, has spectal radius less than one, while $L_1 L_2^M$ has spectral radius larger than one. This is related to the joint spectral radius [16] of the pair of linear operators L_1, L_2 and to the associated finiteness conjecture, see for example [17]. In Section 8 we comment on the relationship between the present work and recent results concerning the joint spectral radius of finite sets of linear operators. Example 3.4 provides an affine IFS on $\mathbb{R}^2$ that has a coding map π, but is not point-fibred on $\mathbb{R}^2$, and hence by

Theorem 1.1, not hyperbolic on $\mathbb{R}^2$. It is, however, point-fibred and hyperbolic when restricted to the x-axis, which is the affine hull of $\pi(\Omega)$, thus illustrating Theorem 1.2.

For the proof of Theorem 1.1 we provide the following roadmap.

(1) The proof that statement (1) $\Rightarrow$ statement (2) is provided in Theorem 4.1.
(2) The proof that statement (2) $\Rightarrow$ statement (3) is provided in Theorem 4.3.
(3) The proof that statement (3) $\Rightarrow$ statement (4) is provided in Theorem 5.10.
(4) The proof that statement (4) $\Rightarrow$ statement (5) is provided in Proposition 6.6.
(5) The proof that statement (5) $\Rightarrow$ statement (1) is provided in Theorem 6.7.

Theorem 1.2 is proved in section 7.

2. Notation and Definitions

We treat $\mathbb{R}^m$ as a vector space, an affine space, and a metric space. We identify a point $x = (x_1, x_2, ..., x_m) \in \mathbb{R}^m$ with the vector whose coordinates are $x_1, x_2, ..., x_m$. We write $0 \in \mathbb{R}^m$ for the point in $\mathbb{R}^m$ whose coordinates are all zero. The standard basis is denoted $\{e_1, e_2, \ldots, e_m\}$. The inner product between $x, y \in \mathbb{R}^m$ is denoted by $\langle x, y \rangle$. The 2-norm of a point $x \in \mathbb{R}^m$ is $\|x\|_2 = \sqrt{\langle x, x \rangle}$, and the euclidean metric $d_E : \mathbb{R}^m \times \mathbb{R}^m \to [0, \infty)$ is defined by $d_E(x, y) = \|x - y\|_2$ for all $x, y \in \mathbb{R}^m$. The following notations, conventions, and definitions will also be used throughout this paper:

(1) A convex body is a compact convex subset of $\mathbb{R}^m$ with non-empty interior.
(2) For a set B in $\mathbb{R}^m$, the notation $conv(B)$ is used to denote the convex hull of B.
(3) For a set $B \in \mathbb{R}^m$, the *affine hull*, denoted $\text{aff}(B)$, of B is the smallest affine subspace containing B, i.e., the intersection of all affine subspaces containing B.
(4) The symbol $\mathbb{H}$ will denote the nonempty compact subsets of $\mathbb{R}^m$, and the symbol $d_{\mathbb{H}}$ will denote the Hausdorff metric on $\mathbb{H}$. Recall that $(\mathbb{R}^m, d_{\mathbb{H}})$ is a complete metric space.

(5) A metric d on $\mathbb{R}^m$ is said to be *Lipschitz equivalent* to d_E if there are positive constants r and R such that

$$r\, d_E(x,y) \leq d(x,y) \leq R\, d_E(x,y),$$

for all $x, y \in \mathbb{R}^m$. If two metrics are Lipschitz equivalent, then they induce the same topology on $\mathbb{R}^m$, but the converse is not necessarily true.

(6) For any two subsets A and B of $\mathbb{R}^m$ the notation $A - B := \{x - y : x \in A \text{ and } y \in B\}$ is used to denote the pointwise subtraction of elements in the two sets.

(7) For a positive integer N, the symbol $\Omega = \{1, 2, \ldots, N\}^\infty$ will denote the set of all infinite sequences of symbols $\{\sigma_k\}_{k=1}^\infty$ belonging to the alphabet $\{1, 2, \ldots, N\}$. The set Ω is endowed with the product topology. An element of $\sigma \in \Omega$ will also be denoted by the concatenation $\sigma = \sigma_1\sigma_2\sigma_3\ldots$, where σ_k denotes the k^{th} component of σ. Recall that since Ω is endowed with the product topology, it is a compact Hausdorff space.

Definition 2.1 (IFS)**.** If $N > 0$ is an integer and $f_n : \mathbb{R}^m \to \mathbb{R}^m$, $n = 1, 2, \ldots, N$, are continuous mappings, then $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is called an *iterated function system* (IFS). If each of the functions in $\mathcal{F}$ is an affine map on $\mathbb{R}^m$, then $\mathcal{F}$ is called an *affine IFS*.

Definition 2.2 (Contractive IFS)**.** An IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is *contractive* when each f_n is a contraction. Namely, there is a number $\alpha_n \in [0, 1)$ such that $d_E(f_n(x), f_n(y)) \leq \alpha_n d_E(x, y)$ for all $x, y \in \mathbb{R}^m$, for all n.

Definition 2.3 (Hyperbolic IFS)**.** An IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is called *hyperbolic* if there is a metric on $\mathbb{R}^m$ Lipschitz equivalent to the given metric so that each f_n is a contraction.

Definition 2.4 (Coding Map)**.** A continuous map $\pi : \Omega \to \mathbb{R}^m$ is called a *coding map* for the IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ if, for each $n = 1, 2, \ldots, N$, the following diagram commutes,

$$\begin{array}{ccc} \Omega & \xrightarrow{s_n} & \Omega \\ \pi \downarrow & & \downarrow \pi \\ \mathbb{R}^m & \xrightarrow[f_n]{} & \mathbb{R}^m \end{array} \tag{2.1}$$

where the symbol $s_n : \Omega \to \Omega$ denotes the inverse shift map defined by $s_n(\sigma) = n\sigma$.

The notion of a coding map is due to J. Kigami [10] and A. Kameyama [8].

Definition 2.5 (Point-Fibred IFS)**.** An IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is *point-fibred* if, for each $\sigma = \sigma_1\, \sigma_2\, \sigma_3 \cdots \in \Omega$, the limit on the right hand side of

$$\pi(\sigma) := \lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x), \tag{2.2}$$

exists, is independent of $x \in \mathbb{R}^m$ for fixed σ, and the map $\pi : \Omega \to \mathbb{R}^m$ is a coding map.

It is not difficult to show that 2.2 is the unique coding map of a point-fibred IFS. Our notion of a point-fibred iterated function system is similar to Kieninger's Definition 4.3.6 [9], p.97. However, we work in the setting of complete metric spaces whereas Kieninger frames his definition in a compact Hausdorff space.

Definition 2.6 (The Symbol $\mathcal{F}(B)$ for an IFS)**.** For an IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ define $\mathcal{F} : \mathbb{H} \to \mathbb{H}$ by

$$\mathcal{F}(B) = \bigcup_{n=1}^{N} f_n(B).$$

(We use the same symbol $\mathcal{F}$ both for the IFS and the mapping.) For $B \in \mathbb{H}$, let $\mathcal{F}^{\circ k}(B)$ denote the k-fold composition of $\mathcal{F}$, i.e., the union of $f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(B)$ over all words $\sigma_1\sigma_2 \cdots \sigma_k$ of length k.

Definition 2.7 (Attractor for an IFS)**.** A set $A \in \mathbb{H}$ is called an *attractor* of an IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ if

$$A = \mathcal{F}(A) \tag{2.3}$$

and

$$A = \lim_{k\to\infty} \mathcal{F}^{\circ k}(B), \tag{2.4}$$

the limit with respect to the Hausdorff metric, for all $B \in \mathbb{H}$.

If an IFS has an attractor A, then clearly A is the unique attractor. It is well known that a hyperbolic IFS has an attractor. An elegant proof of this fact is given by J. Hutchinson [6]. He observes that a contractive IFS $\mathcal{F}$ induces a contraction $\mathcal{F} : \mathbb{H} \to \mathbb{H}$, from

which the result follows by the contraction mapping theorem. See also M. Hata [5] and R. F. Williams [18].

In section 4 it is shown that a point-fibred IFS $\mathcal{F}$ has an attractor A, and, moreover, if π is the coding map of $\mathcal{F}$, then $A = \pi(\Omega)$. Often σ is considered as the "address" of the point $\pi(\sigma)$ in the attractor. In the literature on fractals (for example J. Kigami [10]) there is an approach to the concept of a self-similar system without reference to the ambient space. This approach begins with the idea of a continuous coding map π and, in effect, defines the attractor as $\pi(\Omega)$.

3. Examples and Remarks on Iterated Function Systems

This section contains examples and remarks relevant to Theorems 1.1 and 1.2.

EXAMPLE 3.1 [A Point-fibred, not Contractive IFS] Consider the affine IFS consisting of a single linear function on $\mathbb{R}^2$ given by the matrix

$$f = \begin{pmatrix} 0 & 2 \\ \frac{1}{8} & 0 \end{pmatrix}.$$

Note that the eigenvalues of f equal $\pm\frac{1}{2}$. Since

$$\lim_{n\to\infty} f^{\circ 2n} = \lim_{n\to\infty} T^{-1} \begin{pmatrix} (\frac{1}{2})^n & 0 \\ 0 & (-\frac{1}{2})^n \end{pmatrix} T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

where T is the change of basis matrix, this IFS is point-fibred. However, since

$$f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

the mapping is not a contraction under the usual metric on $\mathbb{R}^2$. Theorem 1.1, however, guarantees we can remetrize $\mathbb{R}^2$ with an equivalent metric so that f is a contraction.

EXAMPLE 3.2 [An IFS with Point-Fibred Functions that is not Point-Fibred] In the literature on affine iterated function systems, it is sometimes assumed that the eigenvalues of the linear parts of the affine functions are less than 1 in modulus. Unfortunately, this assumption is not sufficient to imply any of the five statements given in Theorem 1.1. While the affine IFS $(\mathbb{R}^m; f)$ is point-fibred if and only if the eigenvalues of the linear part of f all have moduli

strictly less than 1, an analogous statement cannot be made if the number of functions in the IFS is larger than 1.

Consider the affine IFS $\mathcal{F} = \left(\mathbb{R}^2; f_1, f_2\right)$, where

$$f_1 = \begin{pmatrix} 0 & 2 \\ \frac{1}{8} & 0 \end{pmatrix} \qquad \text{and} \qquad f_2 = \begin{pmatrix} 0 & \frac{1}{8} \\ 2 & 0 \end{pmatrix}.$$

As noted in Example 3.1

$$\lim_{n\to\infty} f_1^{\circ n}\mathbf{u} = \lim_{n\to\infty} f_2^{\circ n}\mathbf{u} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for any vector $\mathbf{u}$. Thus, both $\mathcal{F}_1 = \left(\mathbb{R}^2; f_1\right)$ and $\mathcal{F}_2 = \left(\mathbb{R}^2; f_2\right)$ are point-fibred. Unfortunately, their product is the matrix

$$f_1 \circ f_2 = \begin{pmatrix} 4 & 0 \\ 0 & \frac{1}{64} \end{pmatrix}, \quad \text{so that} \quad \lim_{n\to\infty} (f_1 \circ f_2)^{\circ n} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lim_{n\to\infty} \begin{pmatrix} 4^n \\ 0 \end{pmatrix} = +\infty.$$

Thus, the IFS $\mathcal{F} = \left(\mathbb{R}^2; f_1, f_2\right)$ fails to be point-fibred.

Remark 3.1.

While it is true that (1) $\Rightarrow$ (2) in Theorem 1.1 even without the assumption that the IFS is affine, the converse is not true in general. Kameyama [8] has shown that there exists a point-fibred IFS that is not hyperbolic. We next give an example of an affine IFS with a coding map that is not point-fibred. Thus, the set of IFSs (with a coding map) strictly contains the set of point-fibred IFSs which, in turn, strictly contains the set of hyperbolic IFSs.

EXAMPLE 3.3 [The Failure of a Finite Eigenvalue Test to Imply Point-Fibred] Consider the linear IFS $\mathcal{F} = \left(\mathbb{R}^2; L_1, L_2\right)$, where

$$L_1 = \begin{pmatrix} 0 & 2 \\ \frac{1}{8} & 0 \end{pmatrix} \qquad \text{and} \qquad L_2 = \begin{pmatrix} a\cos\theta & -a\sin\theta \\ a\sin\theta & a\cos\theta \end{pmatrix} = aR_\theta,$$

where R_θ denotes rotation by angle θ, and $0 < a < 1$. Then L_1^n has eigenvalues $\pm 1/2^n$ while the eigenvalues of L_2^n both have magnitude $a^n < 1$. For example, if we choose $\theta = \pi/8$ and $a = 31/32$ then it is readily verified that the eigenvalues of L_1L_2 and L_2L_1 are smaller than one in magnitude and that one of the eigenvalues of $L_1L_2L_2$ is 1.4014... . Hence, in this case, the magnitudes of the eigenvalues of the linear operators L_1, L_2, L_1^2, L_1L_2, L_2L_1, L_2^2 are all less than one, but $\|(L_1L_2L_2)^n\, x\|$ does not converge when $x \in \mathbb{R}^2$ is any eigenvector of $L_1L_2L_2$ corresponding to the eigenvalue 1.4014.... It

follows that the IFS $\left(\mathbb{R}^2; L_1, L_2\right)$ is not point-fibred. By using the same underlying idea it is straightforward to prove that, given any positive integer M, we can choose a close to 1 and θ close to 0 in such a way that the eigenvalues of $L_{\sigma_1} L_{\sigma_2} ... L_{\sigma_k}$, (where $\sigma_j \in \{1, 2\}$ for $j = 1, 2, ..., k$, with $k \leq M$) are all of magnitude less than one, while $L_1 L_2^M$ has an eigenvalue of magnitude larger than one.

EXAMPLE 3.4 [A non-Hyperbolic Affine IFS] Let $\mathcal{F} = \left(\mathbb{R}^2; f_0, f_1\right)$, where

$$f_0(x_1, x_2) = (\frac{1}{2}\, x_1, x_2), \qquad\qquad f_1(x_1,\, x_2) = (\frac{1}{2}\, x_1 + \frac{1}{2},\, x_2).$$

This IFS has a coding map π with $\Omega = \{0, 1\}^\infty$ and $\pi(\sigma) = (0.\sigma, 0)$, where $0.\sigma$ is considered as a base 2 decimal. Since $\lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x_1, x_2) = (0.\sigma, x_2)$ depends on the choice of the points $(x_1, x_2) \in \mathbb{R}^2$, this IFS cannot be point-fibred. Hence, by Theorem 1.1, the IFS $\mathcal{F}$ is also not hyperbolic. However, it is clearly hyperbolic when restricted to the x-axis, the affine hull of unit interval $\pi(\Omega) = [0, 1] \times \{0\}$. Thus, this example illustrates Theorem 1.2.

A key fact used in the proof of Theorem 1.1 is that the set of antipodal points in a convex body equals the set of diametric points. The definitions of antipodal and diametric points are given in Definitions 6.1 and 6.2, respectively. The equality between these two point sets is proved in Theorem 6.4. While it is possible that this result is present in the convex geometry literature, it does not seem to be well-known. For example, it is not mentioned in the works of Moszynska [13] or Schneider [15]. This equivalence between antipodal and diametric points is crucial to our work because it provides the remetrization technique at the heart of Theorem 6.7, which implies that a non-antipodal IFS is hyperbolic. A consequence of Theorem 1.1 is that a non-antipodal affine IFS has the seemingly stronger property of being topologically contractive.

4. Hyperbolic Implies Point-fibred Implies The Existence of an Attractor

The implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) in Theorem 1.1 are proved in this section. For this section we also introduce the notation $f_{\sigma\,|\,k} = f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x)$. Note that, for k fixed, $f_{\sigma\,|\,k}(x)$ is a function of both x and σ.

Theorem 4.1. *If $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is a hyperbolic IFS, then $\mathcal{F}$ is point fibred.*

Proof. For $\sigma \in \Omega$, the proof that the limit $\lim_{k\to\infty} f_{\sigma|k}$ exists and is independent of x is virtually identical to the proof of the classical Contraction Mapping Theorem. Moreover, the same proof shows that the limit is uniform in σ.

With $\pi : \Omega \to \mathbb{R}^m$ defined by $\pi(\sigma) = \lim_{k\to\infty} f_{\sigma|k}$ it is easy to check that, for each $n = 1, 2, \ldots, N$, the diagram 2.1 commutes.

It only remains to show that π is continuous. With x fixed, $f_{\sigma\,|\,k}(x)$ is a continuous function of σ. This is simply because, if $\sigma, \tau \in \Omega$ are sufficiently close in the product topology, then they agree on the first k components. By Definition 2.5, the function π is then the uniform limit of continuous (in σ) functions defined on the compact set Ω. Therefore, π is continuous. □

Let $\mathcal{F}$ be a point-fibred affine IFS, and let A denote the set

$$A := \pi(\Omega).$$

According to Theorem 4.3, A is the attractor of $\mathcal{F}$.

Lemma 4.2. *Let $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ be a point-fibred affine IFS with coding map $\pi : \Omega \to \mathbb{R}^m$. If $B \subset \mathbb{R}^m$ is compact, then the convergence in the limit*

$$\pi(\sigma) = \lim_{k\to\infty} f_{\sigma|k}(x)$$

is uniform in $\sigma = \sigma_1\sigma_2 \cdots \in \Omega$ and $x \in B$ simultaneously.

Proof. Only the uniformity requiress proof. Express $f_n(x) = L_n x + a_n$, where L_n is the linear part. Then

(4.1)

$$\begin{aligned} f_{\sigma|k}(x) &= L_{\sigma|k}(x) + L_{\sigma|k-1}(a_{\sigma_k}) + L_{\sigma|k-2}(a_{\sigma_{k-1}}) + \cdots + L_{\sigma|1}a_2 + a_1 \\ &= L_{\sigma|k}(x) + f_{\sigma|k}(0). \end{aligned}$$

From equation 4.1 it follows that, for any $x, y \in B$,

(4.2)

$$\begin{aligned} d_E(f_{\sigma|k}(x), f_{\sigma|k}(y)) &= \left\| L_{\sigma|k}(x-y) \right\|_2 \\ &\leq \sup\left\{ \sum_{j=1}^{m} 2\left|c_j\right| \left\| L_{\sigma|k}(e_j) \right\|_2 \,:\, c_1e_1 + \cdots + c_me_m \in B \right\} \\ &\leq c \max_j \left\| f_{\sigma|k}(e_j) - f_{\sigma|k}(0) \right\|_2 , \end{aligned}$$

where $c = 2m \cdot \sup\{\max_j |c_j| \, : c_1 e_1 + \cdots + c_m e_m \in B\}$ and where $\{e_j\}_{j=1}^m$ is a basis for $\mathbb{R}^m$.

Let $\epsilon > 0$. From the definition of point-fibred there is a k_j, independent of σ, such that if $k > k_j$, then

$$\left\| f_{\sigma|k}(e_j) - \pi(\sigma) \right\|_2 < \frac{\epsilon}{4c} \qquad \text{and} \qquad \left\| f_{\sigma|k}(0) - \pi(\sigma) \right\|_2 < \frac{\epsilon}{4c},$$

which implies $\left\| f_{\sigma|k}(e_j) - f_{\sigma|k}(0) \right\|_2 < \frac{\epsilon}{2c}$. This and equation 4.2 implies that if $k \geq \overline{k} := \max_j k_j$, then for any $x, y \in B$ we have

$$d_E(f_{\sigma|k}(x), f_{\sigma|k}(y)) < c\frac{\epsilon}{2c} = \frac{\epsilon}{2}. \tag{4.3}$$

Let b be a fixed element of B. There is a k_b, independent of σ, such that if $k > k_b$, then $d_E(f_{\sigma|k}(b), \pi(\sigma)) < \frac{\epsilon}{2}$. If $k > max(k_b, \overline{k})$ then, by equation 4.3, for any $x \in B$

$$d_E(f_{\sigma|k}(x), \pi(\sigma)) \leq d_E(f_{\sigma|k}(x), f_{\sigma|k}(b)) + d_E(f_{\sigma|k}(b), \pi(\sigma)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Theorem 4.3 (A Point-Fibred IFS has an Attractor)**.** *If $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is a point-fibred affine IFS, then $\mathcal{F}$ has an attractor $A = \pi(\Omega)$, where $\pi : \Omega \to \mathbb{R}^m$ is the coding map of $\mathcal{F}$.*

Proof. It follows directly from the commutative diagram (2.1) that A obeys the self-referential equation (2.3). We next show that A satisfies equation (2.4).

Let $\epsilon > 0$. We must show that there is an M such that if $k > M$, then $d_{\mathbb{H}}(\mathcal{F}^{\circ k}(B), \pi(\Omega)) < \epsilon$. It is sufficient to let $M = \max(M_1, M_2)$, where M_1 and M_2 are defined as follows.

First, let a be an arbitrary element of A. Then there exists a $\sigma \in \Omega$ such that $a = \pi(\sigma)$. By Lemma 4.2 there is an M_1 such that if $k > M_1$, then $d_E(f_{\sigma|k}(b), a) = d_E(f_{\sigma|k}(b), \pi(\sigma)) < \epsilon$, for all $b \in B$. In other words, A lies in an ϵ-neighborhood of $\mathcal{F}^{\circ k}(B)$.

Second, let b be an arbitrary element of B and σ an arbitrary element of Ω. If $a := \pi(\sigma) \in A$, then there is an M_2 such that if $k > M_2$, then $d_E(f_{\sigma|k}(b), a) = d_E(f_{\sigma|k}(b), \pi(\sigma)) < \epsilon$. In other words, $\mathcal{F}^{\circ k}(B)$ lies in an ϵ-neighborhood of A. □

5. An IFS with an Attractor is Topologically Contractive

The goal of this section is to establish the implication (3) $\Rightarrow$ (4) in Theorem 1.1. We will show that if an affine IFS has an attractor as defined in Defintion 2.7 , then it is a topological contraction. The proof uses notions involving convex bodies.

Definition 5.1. A convex body K is *centrally symmetric* if it has the property that whenever $x \in K$, then $-x \in K$.

A well-known general technique for creating centrally symmetric convex bodies from a given convex body is provided by the next proposition.

Proposition 5.2. *If a set K is a convex body in $\mathbb{R}^m$, then the set $K' = K - K$ is a centrally symmetric convex body in $\mathbb{R}^m$.*

Definition 5.3 (Minkowski Norm)**.** If K is a centrally symmetric convex body in $\mathbb{R}^m$, then the *Minkowski norm* on $\mathbb{R}^m$ is defined by

$$\|x\|_K = \inf\{\lambda \geq 0 \,:\, x \in \lambda K\}.$$

The next proposition is also well-known.

Proposition 5.4. *If K is a centrally symmetric convex body in $\mathbb{R}^m$, then the function $\|x\|_K$ defines a norm on $\mathbb{R}^m$. Moreover, the set K is the unit ball with respect to the Minkowski norm $\|x\|_K$.*

Definition 5.5 (Minkowski Metric)**.** If K is a centrally symmetric convex body in $\mathbb{R}^m$ and $\|x\|_K$ is the associated Minkowski norm, then define the *Minkowski metric* on $\mathbb{R}^m$ by the rule

$$d_K(x, y) := \|x - y\|_K \, .$$

While R. Rockafeller [14] refers to such a metric as a *Minkowski metric*, the reader should be aware that this term is also associated with the metric on space-time in theory of relativity. Since, for any convex body K there are positive numbers r and R such that K contains a ball of radius r and is contained in a ball of radius R, the following proposition is clear.

Proposition 5.6. *If d is a Minkowski metric, then d is Lipschitz equivalent to the standard metric d_E on $\mathbb{R}^m$.*

Proposition 5.7. *A metric $d : \mathbb{R}^m \times \mathbb{R}^m \to [0,\infty)$ is a Minkowski metric if and only if it is translation invariant and distances behave linearly along line segments. More specifically,*

(5.1)
$$d(x+z, y+z) = d(x,y) \qquad \text{and} \qquad d(x,(1-\lambda)x+\lambda y) = \lambda d(x,y)$$

for all $x, y, z \in \mathbb{R}^m$ and all $\lambda \in [0,1]$.

Proof. For a proof see Rockafeller [14] pp.131-132. □

Definition 5.8 (Topologically Contractive IFS)**.** An IFS $\mathcal{F} = \{\mathbb{R}^m; f_1, f_2, ..., f_N\}$ is called *topologically contractive* if there is a convex body K such that $\mathcal{F}(K) \subset int(K)$.

The proof of Theorem 5.10 relies on the following lemma which is easily proved.

Lemma 5.9. *If $g : \mathbb{R}^m \to \mathbb{R}^m$ is affine and $S \subset \mathbb{R}^m$, then $g(conv(S)) = conv(g(S))$.*

Theorem 5.10 (The Existence of an Attractor Implies a Topological Contraction)**.** *For an affine IFS $\mathcal{F} = \{\mathbb{R}^m; f_1, f_2, ..., f_N\}$, if there exists an attractor $A \in \mathbb{H}$ of the affine IFS $\mathcal{F} = \{\mathbb{R}^m; f_1, f_2, ..., f_N\}$, then $\mathcal{F}$ is topologically contractive.*

Proof. The proof of this theorem unfolds in three steps.

(1) There exists a convex body K_1 and a positive integer t with the property that $\mathcal{F}^{\circ t}(K_1) \subset int(K_1)$.
(2) The set K_1 is used to define a convex body K_2 such that $L_n(K_2) \subset int(K_2)$, where $f_n(x) = L_n x + a_n$ and $n = 1, 2, \ldots, N$.
(3) There is a positive constant c such that the set $K = cK_2$ has the property $\mathcal{F}(K) \subset int(K)$.

Proof of Step (1). Let A denote the attractor of $\mathcal{F}$. Let $A_\rho = \{x \in \mathbb{R}^m : d_{\mathbb{H}}(\{x\}, A) \le \rho\}$ denote the dilation of A by radius $\rho > 0$. Since we are assuming $\lim_{k\to\infty} d_{\mathbb{H}}(\mathcal{F}^{\circ k}(A_\rho), A) = 0$, we can find an integer t so that $d_{\mathbb{H}}(\mathcal{F}^{\circ t}(A_1), A) < 1$. Thus,

$$\mathcal{F}^{\circ t}(A_1) \subseteq int(A_1). \tag{5.2}$$

If we let $K_1 := conv\,(A_1)$, then

$$\begin{aligned}
\mathcal{F}^{\circ t}(K_1) &= \bigcup_{i_1\in\Omega}\bigcup_{i_2\in\Omega}\cdots\bigcup_{i_t\in\Omega}(f_{i_1}\circ f_{i_2}\circ\cdots\circ f_{i_t})\,(conv(A_1))\\
&= \bigcup_{i_1\in\Omega}\bigcup_{i_2\in\Omega}\cdots\bigcup_{i_t\in\Omega} conv\,(f_{i_1}\circ f_{i_2}\circ\cdots\circ f_{i_t}(A_1)) \qquad \text{(by Lemma 5.9)}\\
&\subseteq \bigcup_{i_1\in\Omega}\bigcup_{i_2\in\Omega}\cdots\bigcup_{i_t\in\Omega} conv\,(int(A_1)) = conv(int(A_1)) \quad \text{(by inclusion (5.2))}\\
&\subseteq int(conv(A_1)) = int\,(K_1)\,.
\end{aligned}$$

This argument completes the proof of Step (1).

Proof of Step (2). Consider the set

$$K_2 := \sum_{k=0}^{t-1}(conv(\mathcal{F}^{\circ k}(K_1) - conv(\mathcal{F}^{\circ k}(K_1))).$$

The set K_2 is a centrally symmetric convex body because it is a finite Minkowski sum of centrally symmetric convex bodies. If any affine map f_n in $\mathcal{F}$ is written $f_n(x) = L_n x + a_n$, where $L_n : \mathbb{R}^m \to \mathbb{R}^m$ denotes the linear part, then

$$\begin{aligned}
L_n(K_2) &= \sum_{k=0}^{t-1} L_n\Big(conv(\mathcal{F}^{\circ k}(K_1) - conv(\mathcal{F}^{\circ k}(K_1))\Big) \quad \text{(since } L_n \text{ is a linear map)}\\
&= \sum_{k=0}^{t-1}\Big(conv(L_n\left(\mathcal{F}^{\circ k}(K_1)\right)) - conv(L_n\left(\mathcal{F}^{\circ k}(K_1)\right))\Big) \quad \text{(by Lemma 5.9)}\\
&= \sum_{k=0}^{t-1}\Big(conv(f_n\left(\mathcal{F}^{\circ k}(K_1)\right)) - conv(f_n\left(\mathcal{F}^{\circ k}(K_1)\right))\Big) \quad \text{(since the } a_n\text{s cancel)}\\
&\subseteq \sum_{k=0}^{t-1}\Big(conv(\mathcal{F}^{\circ(k+1)}(K_1)) - conv(\mathcal{F}^{\circ(k+1)}(K_1))\Big)\\
&= \Big(conv(\mathcal{F}^{\circ t}(K_1) - conv(\mathcal{F}^{\circ t}(K_1))\Big) + \sum_{k=1}^{t-1}\Big(conv(\mathcal{F}^{\circ k}(K_1) - conv(\mathcal{F}^{\circ k}(K_1))\Big)\\
&\subseteq (int\,(K_1) - int\,(K_1)) + \sum_{k=1}^{t-1}(conv(\mathcal{F}^{\circ k}(K_1) - conv(\mathcal{F}^{\circ k}(K_1))) \quad \text{(by Step 1)}\\
&= int(K_2).
\end{aligned}$$

The second to last inclusion follows from the fact that $f_n\left(\mathcal{F}^{\circ k}\left(K_1\right)\right) \subset \mathcal{F}^{\circ(k+1)}\left(K_1\right)$. The last equality follows from the fact that if $\mathcal{O}$ and $\mathcal{C}$ are symmetric convex bodies in $\mathbb{R}^m$, then $int(\mathcal{O}) + \mathcal{C} = int\left(\mathcal{O} + \mathcal{C}\right)$. We have now completed the proof of Step (2).

Proof of Step (3). It follows from Step (2) and the compactness of K_2 that there is a constant $\alpha \in (0,1)$ such that $d_{K_2}(L_n(x), L_n(y)) < \alpha d_{K_2}(x,y)$ for all $x, y \in \mathbb{R}^m$ and all $n = 1, 2, \ldots, N$.

Let
$$c > \frac{r}{(1-\alpha)},$$
where $r = \max\{d_{K_2}(a_1, 0), d_{K_2}(a_2, 0), \ldots, d_{K_2}(a_N, 0)\}$. If $x \in cK_2$ and $f(x) = Lx + a$ is any function in the IFS $\mathcal{F}$, then

$$\begin{aligned}
\|f(x)\|_{K_2} &= d_{K_2}\left(f\left(x\right), 0\right) = d_{K_2}\left(Lx + a, 0\right) \leq d_{K_2}\left(Lx + a, Lx\right) + d_{K_2}\left(Lx, 0\right) \\
&= d_{K_2}\left(a, 0\right) + d_{K_2}\left(Lx, 0\right) \quad \text{(by Equation (5.1))} \\
&< r + \alpha d_{K_2}\left(x, 0\right) = r + \alpha \left\|x\right\|_{K_2} \\
&\leq r + \alpha c < (c - \alpha c) + \alpha c = c.
\end{aligned}$$

This inequality shows that $\mathcal{F}\left(cK_2\right) \subset int(cK_2)$. □

6. A Non-Antipodal Affine IFS is Hyperbolic

Let $S^{m-1} \subset \mathbb{R}^m$ denote the unit sphere in $\mathbb{R}^m$. For a convex body $K \subset \mathbb{R}^m$ and $u \in S^{m-1}$ there exists a pair, $\{H_u, H_{-u}\}$, of distinct supporting hyperplanes of K, each orthogonal to u and with the property that they both intersect ∂K but contain no points of the interior of K. Since by definition a convex body has non-empty interior, this pair will be unique. The pair $\{H_u, H_{-u}\}$ is usually referred to as the two *supporting hyperplanes* of K orthogonal to u. (See Moszynska [13], p.14.)

Definition 6.1 (Antipodal Pairs)**.** If $K \subset \mathbb{R}^m$ is a convex body and $u \in S^{m-1}$, then define

$$\mathcal{A}_u := \mathcal{A}_u(K) = \{(p,q) \in (H_u \cap \partial K) \times (H_{-u} \cap \partial K)\} \quad \text{and}$$
$$\mathcal{A} := \mathcal{A}(K) = \bigcup_{u \in S^{m-1}} \mathcal{A}_u.$$

We say that (p,q) is an *antipodal pair* of points with respect to K if $(p,q) \in \mathcal{A}$.

Definition 6.2 (Diametric Pairs)**.** If $K \subset \mathbb{R}^m$ is a convex body, and $u \in S^{m-1}$, then define the *diameter* of K *in the direction* u to be

$$D(u) = \max\{\|x-y\|_2 : x, y \in K, x-y = \alpha u, \alpha \in \mathbb{R}\}.$$

The maximum is achieved at some pair of points belonging to ∂K because $K \times K$ is convex and compact, and $\|x-y\|_2$ is continuous for $(x,y) \in K \times K$. Now define

$$\mathcal{D}_u = \{(p,q) \in \partial K \times \partial K : D(u) = \|q-p\|_2\} \quad \text{and}$$
$$\mathcal{D} = \bigcup_{u \in S^{m-1}} \mathcal{D}_u.$$

We say that $(p,q) \in \mathcal{D}_u$ is a *diametric pair* of points in the direction of u, and that $\mathcal{D}$ is the set of diametric pairs of points of K.

Definition 6.3 (Strictly Convex)**.** A convex body K is *strictly convex* if, for every two points $x, y \in K$, the open line segment joining x and y is contained in the interior of K.

We write xy to denote the closed line segment with endpoints at x and y so that $y-x$ is the vector, in the direction from x to y, whose magnitude is the length of xy.

Theorem 6.4. *If $K \subset \mathbb{R}^m$ is a convex body, then the set of antipodal pairs of points of K is the same as the set of diametric pairs of points of K, i.e.,*

$$\mathcal{A} = \mathcal{D}.$$

Proof. First we show that $\mathcal{A} \subseteq \mathcal{D}$. If $(p,q) \in \mathcal{A}$, then $p \in H_u \cap \partial K$ and $q \in H_{-u} \cap \partial K$ for some $u \in S^{m-1}$. Clearly any chord of K parallel to pq lies entirely in the region between H_u and H_{-u} and therefore cannot have length greater than that of pq. So $D(q-p) = \|q-p\|$ and $(p,q) \in \mathcal{D}_{q-p} \subseteq \mathcal{D}$. Note, for use later in the proof, that if K is strictly convex, then pq is the unique chord of maximum length in its direction.

Conversely, to show that $\mathcal{D} \subseteq \mathcal{A}$, first consider the case where K is a strictly convex body. For each $u \in S^{m-1}$, consider the points $x_u \in H_u \cap \partial K$ and $x_{-u} \in H_{-u} \cap \partial K$. The continuous function $f : S^{m-1} \to S^{m-1}$ defined by

$$f(u) = \frac{x_u - x_{-u}}{\|x_u - x_{-u}\|_2}$$

has the property that $\langle f(u), u\rangle > 0$ for all u. In other words, the angle between u and $f(u)$ is less than $\frac{\pi}{2}$. But it is an elementary exercise in topology (see, for example, Munkres [12], problem 10, page 367) that if $f : S^{m-1} \to S^{m-1}$ maps no point x to its antipode $-x$, then f has degree 1 and, in particular, is surjective. To show that $\mathcal{D} \subseteq \mathcal{A}$, let $(p, q) \in \mathcal{D}_v$ for some $v \in S^{m-1}$. By the surjectivity of f there is $u \in S^{m-1}$ such that $f(u) = v$. According to the last sentence of the previous paragraph, $x_u x_{-u}$ is the unique longest chord parallel to v. Therefore $p = x_u$ and $q = x_{-u}$ and consequently $(p, q) \in \mathcal{A}_u$.

The case where K is not strictly convex is treated by a standard limiting argument. Given a vector $v \in S^{m-1}$ and a longest chord pq parallel to v, we must prove that $(p, q) \in \mathcal{A}$. Since K is the intersection of all strictly convex bodies containing K, there is a sequence $\{K_k\}$ of strictly convex bodies containing K with the following two properties.

1. There is a longest chord $p_k q_k$ of K_k parallel to u such that $\lim_{k\to\infty} \|p_k - q_k\|_2 = \|p - q\|_2$, and the limits $\lim_{k\to\infty} p_k = \tilde{p} \in K$ and $\lim_{k\to\infty} q_k = \tilde{q} \in K$ exist.

By the result for the strictly convex case, there is a sequence of vectors $u_k \in S^{m-1}$ such that $p_k = K_k \cap H_{u_k}(K_k)$ and $q_k = K_k \cap H_{-u_k}(K_k)$. By perhaps going to a subsequence

2. $\lim_{k\to\infty} u_k = u \in S^{m-1}$ exists.

It follows from item 1 that $\|\tilde{p} - \tilde{q}\|_2 = \|p - q\|_2$ and $\tilde{p} - \tilde{q}$ is parallel to v. Therefore, $\tilde{p}\tilde{q}$ as well as pq, are longest chords of K parallel to v. It follows from 2 that if H and H' are the hyperplanes orthogonal to u through $\tilde{p}$ and $\tilde{q}$ respectively, then H and H' are parallel supporting hyperplanes of K. Therefore, necessarily $p \in H$ and $q \in H'$, and consequently $(p, q) \in \mathcal{A}_u \subset \mathcal{A}$. $\square$

Definition 6.5 (Non-Antipodal IFS)**.** If $K \subset \mathbb{R}^m$ is a convex body, then $f : \mathbb{R}^m \to \mathbb{R}^m$ is *non-antipodal* with respect to K if $f(K) \subseteq K$ and $(x, y) \in \mathcal{A}(K)$ implies $(f(x), f(y)) \notin \mathcal{A}(K)$. If $\mathcal{F} = \{\mathbb{R}^m; f_1, f_2, ..., f_N\}$ is an iterated function system with the property that each f_n is non-antipodal with respect to K, then $\mathcal{F}$ is called *non-antipodal* with respect to K.

The next proposition gives the implication (4) $\Rightarrow$ (5) in Theorem 1.1. The proof is clear.

Proposition 6.6 (A Topological Contraction is Non-Antipodal)**.** *If $\mathcal{F} = \{\mathbb{R}^m; f_1, f_2, ..., f_N\}$ is an affine iterated function system with the property that there exists a convex body $K \subset \mathbb{R}^m$ such that $f_n(K) \subset int(K)$ for all $n = 1, 2, \ldots, n$, then $\mathcal{F}$ is non-antipodal with respect to K.*

The next theorem provides the implication that (5) $\Rightarrow$ (1) in Theorem 1.1.

Theorem 6.7. *If the affine IFS $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is non-antipodal with respect to a convex body K, then $\mathcal{F}$ is hyperbolic.*

Proof. Assume that K is a convex body such that f is non-antipodal with respect to K for all $f \in \mathcal{F}$. Let $C = K - K$ and let $f(x) = Lx + a \in \mathcal{F}$, where L is the linear part of f. By Proposition 5.2, the set C is a centrally symmetric convex body and

$$L(C) = L(K) - L(K) = f(K) - f(K) \subseteq K - K = C.$$

We claim that $L(C) \subset int(C)$. Since C is compact and L is linear, to prove the claim it is sufficient to show that $L(x) \notin \partial C$ for all $x \in \partial C$. By way of contradiction, assume that $x \in \partial C$ and $L(x) \in \partial C$. Then the vector x is a longest vector in C in its direction. Since $x \in C = K - K$ there are $x_1, x_2 \in \partial K$ such that $x = x_1 - x_2$, and $(x_1, x_2) \in \mathcal{D}(K) = \mathcal{A}(K)$, where the last equality is by Theorem 6.4. So (x_1, x_2) is an antipodal pair with respect to K. Likewise, since Lx is a longest vector in C in its direction, there are $y_1, y_2 \in \partial K$ such that $Lx = y_1 - y_2$, and $(y_1, y_2) \in \mathcal{D}(K) = \mathcal{A}(K)$. Therefore

$$f(x_2) - f(x_1) = L(x_2) - L(x_1) = L(x_2 - x_1) = Lx = y_1 - y_2,$$

which implies that $(f_n(x_1), f_n(x_2)) \in \mathcal{D}(K) = \mathcal{A}(K)$, contradicting that f is non-antipodal with respect to K.

If d_C denotes the Minkowski metric with respect to the centrally symmetric convex body C, then by Proposition 5.4 C is the unit ball centered at the origin with respect to this metric. Since C is compact, the containment $L(C) \subset int(C)$ implies that there is an $\alpha \in [0, 1)$ such that $\|Lx\|_C < \alpha \; \|x\|_C$ for all $x \in \mathbb{R}^m$. Then

$$\begin{aligned} d_C(f(x), f(y)) &= \|f(x) - f(y)\|_C = \|Lx - Ly\|_C \\ &= \|L(x - y)\|_C < \alpha \; \|x - y\|_C = \alpha \, d_C(x, y). \end{aligned}$$

Therefore d_C is a metric for which each function in the IFS is a contraction. By Proposition 5.6, d_C is Lipschitz equivalent to the standard metric. □

7. An Answer to the Question of Kameyama

We now turn to the proof of Theorem 1.2, the theorem that settles the question of Kameyama. If $X \subseteq \mathbb{R}^m$ and $\mathcal{F} = (X; f_1, f_2, ..., f_N)$ is an IFS on X, then the definitions of *coding map* and *point-fibred* for $\mathcal{F}$ are exactly the same as Definitions 2.4 and 2.5, with $\mathbb{R}^m$ replaced by X. The proof of Theorem 1.2 requires the following proposition.

Proposition 7.1. *If $X \subseteq \mathbb{R}^m$ and $\mathcal{F} = (X; f_1, f_2, ..., f_N)$ is an IFS with a coding map $\pi : \Omega \to \mathbb{R}^m$ such that $\pi(\Omega) = X$, then $\mathcal{F}$ is point-fibred on X.*

Proof. By Definition 2.5, we must show that $\lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x)$ exists, is independent of $x \in X$, and is continuous as a function of $\sigma = \sigma_1\sigma_2 \cdots \in \Omega$. We will actually show that $\lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x) = \pi(\sigma)$.

Since π is a coding map, we know by Definition 2.4 that $f_n \circ \pi(\sigma) = \pi \circ s_n(\sigma)$, for all $n = 1, 2, \ldots, N$. By assumption, if x is any point in X, then there is a $\tau \in \Omega$ such that $\pi(\tau) = x$. Thus

$$\begin{aligned}
\lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x) &= \lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(\pi(\tau)) \text{ (since } \pi(\tau) = x) \\
&= \lim_{k\to\infty} \pi(s_{\sigma_1} \circ s_{\sigma_2} \circ ... \circ s_{\sigma_k} \circ \tau) \text{ (by Diagram 2.1)} \\
&= \pi(\lim_{k\to\infty} s_{\sigma_1} \circ s_{\sigma_2} \circ ... \circ s_{\sigma_k} \circ \tau) \text{ (since } \pi \text{ is continuous)} \\
&= \pi(\sigma).
\end{aligned}$$

□

Theorem 7.2. *If $\mathcal{F} = (\mathbb{R}^m; f_1, f_2, ..., f_N)$ is an affine IFS with a coding map $\pi : \Omega \to X$, then $\mathcal{F}$ is point-fibred when restricted to the affine hull of $\pi(\Omega)$. In particular, if $\pi(\Omega)$ contains a non-empty open subset of $\mathbb{R}^m$, then $\mathcal{F}$ is point-fibred on $\mathbb{R}^m$.*

Proof. Let $A := \pi(\Omega)$. Since $f_n(A) \subseteq A$ for all n, the restriction of the IFS $\mathcal{F}$ to A, namely $\mathcal{F}|_A := (A; f_1, f_2, \ldots, f_N)$, is well defined. It follows from Proposition 7.1 that $\mathcal{F}|_A$ is point-fibred and, because

the coding map for a point-fibred IFS is unique,

$$\pi(\sigma) = \lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(a)$$

for $(\sigma, a) \in \Omega \times A$. It only remains to show that the restriction $\mathcal{F}|_{\mathrm{aff}(A)} := (\mathrm{aff}(A); f_1, f_2, \ldots, f_N)$ of the affine IFS $\mathcal{F}$ to the affine hull of A is point-fibred.

Let $x \in \mathrm{aff}(A)$, the affine hull of A. It is well known that any point in the affine hull can be expressed as a sum, $x = \sum_{p=0}^{m} \lambda_p a_p$ for some $\lambda_0, \lambda_1, ..., \lambda_m \in \mathbb{R}$ such that $\sum_{p=0}^{m} \lambda_p = 1$ and $a_0, a_1, ..., a_m \in A$. Hence, for $(\sigma, x) \in \Omega \times \mathrm{aff}(A)$,

$$\begin{aligned}
\lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x) &= \lim_{k\to\infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(\sum_{p=0}^{m} \lambda_p a_p), \\
&= \lim_{k\to\infty} \sum_{p=0}^{m} \lambda_p f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(a_p) \\
&= \sum_{p=0}^{m} \lambda_p \pi(\sigma) = \pi(\sigma).
\end{aligned}$$

□

Theorem 1.2 now follows easily from Theorem 7.2 and Theorem 1.1.

Proof. (of Theorem 1.2) Let $A := \pi(\Omega)$ and let $\dim \mathrm{aff}(A) = k \leq m$. It is easy to check from the commuting diagram 2.1 that $f(A) \subseteq A$ for each $f \in \mathcal{F}$ implies that $f(\mathrm{aff}(A)) \subseteq \mathrm{aff}(A)$ for each $f \in \mathcal{F}$. Since $\mathrm{aff}(A)$ is isomorphic to $\mathbb{R}^k$, Theorem 1.1 can be applied to the IFS $\mathcal{F}|_{\mathrm{aff}(A)} := (\mathrm{aff}(A); f_1, f_2, ...f_N)$ to conclude that, since it is point-fibred, $\mathcal{F}|_{\mathrm{aff}(A)}$ is also hyperbolic. □

Note that the IFS $(\mathbb{R}; f)$, where $f(x) = 2x + 1$, is not hyperbolic on $\mathbb{R}$, but it is hyperbolic on the affine subspace $\{-1\} \subset \mathbb{R}$.

8. Concluding Remarks

Recently it has come to our attention that another condition, equivalent to conditions (1) – (5) in our main result, Theorem 1.1, is (6) $\mathcal{F}$ *has joint spectral radius less than one.* (We define the joint spectral radius (JSR) of an affine IFS to be the joint spectral radius of the set of linear factors of its maps.) This information is

important because it connects our approach to the rapidly growing literature about JSR, see for example [3], [4], and works that refer to these.

Since Example 3.3 and the results presented by Blondel, Theys, and Vladimirov [17] indicate there is no general fast algorithm which will determine whether or not the joint spectral radius of an IFS is less than one, we feel that Theorem 1.1 is important because it provides an easily testable condition that an IFS has a unique attractor. In particular, the topologically contractive and non-antipodal conditions (conditions 4 and 5) provide geometric/visual tests, which can easily be checked for any affine IFS. In addition to yielding the existence of an attractor, these two conditions also provide information concerning the location of the attractor. (For example, the attractor is a subset of a particular convex body.) We also anticipate that Theorem 1.1 can be generalized into other broader classes of functions, where the techniques developed for the theory of joint spectal radius will not apply.

References

[1] M. F. Barnsley, *Fractal image compression,* Notices Am. Math. Soc. **43** (1996) 657-662.

[2] M. F. Barnsley, V. Ervin, D. Hardin, and J. Lancaster, *Solution of an inverse problem for fractals and other sets,* Proc. Nat. Acad. Sci. **83** (1985) 1975-1977.

[3] M. A. Berger and Y. Wang, *Bounded semigroups of matrices*, Linear Algebra and Appl. **166** (1992), 21-27.

[4] I. Daubechies and J.C. Lagarias, *Sets of matrices all infinite products of which converge*, Linear Algebra and Appl. **162** (1992), 227-263.

[5] M. Hata, *On the structure of self-similar sets,* Japan J. Appl. Math. **2** (1985) 381–414.

[6] J. E. Hutchinson, *Fractals and self-similarity,* Indiana Univ. Math. J. **30** (1981) 713–747.

[7] L. Janos, *A converse of Banach's contraction theorem,* Proc. Amer. Math. Soc. **18** (1967) 287–289.

[8] Atsushi Kameyama, *Distances on Topological Self-Similar Sets,* Proceedings of Symposia in Pure Mathematics, Volume **72.1,** 2004.

[9] Bernd Kieninger, *Iterated Function Systems on Compact Hausdorff Spaces,* Shaker Verlag, Aachen, 2002.

[10] J. Kigami, *Analysis on Fractals*, Cambridge University Press, 2001.

[11] S. Leader, *A topological characterization of Banach contractions,* Pac. Jour. Math. **69** (1977) 461–466.

[12] James R. Munkres, *Topology* (second edition), Prentice Hall, Upper Saddle River, NU, 2000.

[13] Maria Moszyńska, *Selected Topics in Convex Geometry,* Birkhäuser, Boston, 2006.

[14] R. Tyrrel Rockafeller, *Convex Analysis,* Princeton University Press, Princeton, 1970.

[15] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory,* Cambridge University Press, 1993.

[16] Gian-Carlo Rota and W. Gilbert Strang, *A note on the joint spectral radius,* Nederl. Akad. Wetensch. Proc. Ser. A 63=Indag. Math. **22** (1960) 379-381.

[17] Vincent Blondel, Jacques Theys, and Alexander Vladimirov, *An elementary counterexample to the finiteness conjecture,* SIAM J. on Matrix Anal. and Appl. **24** (2003) 963-970.

[18] R. F. Williams, *Composition of contractions,* Bol. da Soc. Brazil de Mat. **2** (1971) 55-59.

[19] Roger Webster, *Convexity*, Oxford University Press, Oxford, 1994.

Department of Mathematics, Australian National University, Canberra, ACT, Australia

E-mail address: michael.barnsley@maths.anu.edu.au, mbarnsley@aol.com

URL: http://www.superfractals.com

Department of Mathematics, University of Florida, Gainesville, FL 32611-8105, USA

E-mail address: avince@math.ufl.edu

URL: http://www.math.ufl.edu/∼vince/

Department of Mathematics, University of Florida, Gainesville, FL 32611-8105, USA

E-mail address: dcw@math.ufl.edu

URL: http://www.math.ufl.edu/∼dcw/